

UNIVERSIDAD COMPLUTENSE DE MADRID

ESCUELA UNIVERSITARIA DE ÓPTICA
Departamento de Óptica II (Optometría y Visión)



**FIABILIDAD INTRAEXAMINADOR Y
CONCORDANCIA DE PRUEBAS CLÍNICAS DE
EVALUACIÓN DE LA VISIÓN BINOCULAR.**

**MEMORIA PARA OPTAR AL GRADO DE DOCTOR
PRESENTADA POR**

Beatriz Antona Peñalba

Bajo la dirección del doctor

Francisco Barra Lázaro

Madrid, 2010

- ISBN: 978-84-693-0657-4

UNIVERSIDAD COMPLUTENSE DE MADRID
DEPARTAMENTO DE ÓPTICA II (OPTOMETRÍA Y VISIÓN)



TESIS DOCTORAL

FIABILIDAD INTRAEXAMINADOR Y CONCORDANCIA
DE PRUEBAS CLÍNICAS DE EVALUACIÓN DE LA
VISIÓN BINOCULAR

BEATRIZ ANTONA PEÑALBA

MADRID, 2008

Departamento de Óptica II (Optometría y Visión)

UNIVERSIDAD COMPLUTENSE DE MADRID



FIABILIDAD INTRAEXAMINADOR Y CONCORDANCIA DE PRUEBAS CLÍNICAS DE EVALUACIÓN DE LA VISIÓN BINOCULAR

BEATRIZ ANTONA PEÑALBA

Director: Dr. FRANCISCO BARRA LÁZARO

Agradecimientos

Quiero aprovechar estas líneas para expresar mi más sincero agradecimiento a las siguientes personas:

A Dra. Ana Rosa Barrio, por su inestimable ayuda, sugerencias aportadas y su permanente disposición.

A mis compañeros: Enrique González e Isabel Sánchez; así como a José Luis Cebrian y a Guadalupe González por su ayuda en la toma de medidas.

A Alberto, por animarme en mis decisiones y su importante apoyo en los momentos difíciles.

También quiero dar las gracias a todos los alumnos que han querido colaborar voluntaria y desinteresadamente en este estudio por su paciencia e interés mostrados.

Y por supuesto agradezco a mi director de tesis Dr. Francisco Barra sus consejos y experiencia prestada que me han ayudado mucho para realizar la presente investigación.

A Alberto por su apoyo y por ser mi alegría de cada día y
a mis padres, Estefanía y Agapito, por sus enseñanzas y por
su cariño desinteresado

“Son muchas las preguntas que pueden plantearse simplemente observando la realidad que nos rodea. (...) Por lo general, las respuestas no son el final del camino. Por el contrario, a menudo abren un inmenso abanico de nuevos, y más complejos, interrogantes. El proceso así engendrado no conoce límites. (...) El trayecto se seguirá haciendo observando, preguntando, analizando y respondiendo.

(Fragmento extraído del panel de presentación del Museo Nacional de la Ciencia y la Tecnología de Madrid, DIC 2006)

ÍNDICE

<i>Lista de símbolos, abreviaturas y siglas.....</i>	<i>iii</i>
<i>Glosario de términos usados en estudios de medición.....</i>	<i>v</i>
<i>Lista de tablas.....</i>	<i>vii</i>
<i>Lista de figuras.....</i>	<i>ix</i>

PARTE I. ASPECTOS GENERALES

1. INTRODUCCIÓN.....	1
1.1. OBJETIVOS DE ESTA INVESTIGACIÓN.....	3
1.2. APORTACIONES NOVEDOSAS	5
1.3. ESTRUCTURA DE LA TESIS.....	14
2. FUNDAMENTACIÓN TEÓRICA	17
2.1. ESTUDIOS DE REPETIBILIDAD DE PRUEBAS CLÍNICAS	21
2.2. ESTUDIOS DE CONCORDANCIA DE PRUEBAS CLÍNICAS.....	25
3. METODOLOGÍA GENERAL	27
3.1. SUJETOS	27
3.2. PROCEDIMIENTO	32
3.3. TRATAMIENTO ESTADÍSTICO DE LOS DATOS.....	36

PARTE II. ESTUDIO EMPÍRICO

4. AMPLITUD DE ACOMODACIÓN	39
4.1. INTRODUCCIÓN	39
4.2. MÉTODO.....	41
4.3. RESULTADOS	43
4.4. DISCUSIÓN.....	48
4.5. CONCLUSIONES	52
5. FLEXIBILIDAD ACOMODATIVA	55
5.1. INTRODUCCIÓN	55
5.2. MÉTODO.....	57
5.3. RESULTADOS	58
5.4. DISCUSIÓN.....	60
5.5. CONCLUSIONES	63
6. RESPUESTA ACOMODATIVA.....	65
6.1. INTRODUCCIÓN	65
6.2. MÉTODO.....	67
6.3. RESULTADOS	73
6.4. DISCUSIÓN.....	78
6.5. CONCLUSIONES	82

7.	ACOMODACIÓN RELATIVA	83
7.1.	INTRODUCCIÓN	83
7.2.	MÉTODO	83
7.3.	RESULTADOS	84
7.4.	DISCUSIÓN.....	86
7.5.	CONCLUSIONES	86
8.	DESVIACIÓN OCULAR	87
8.1.	INTRODUCCIÓN	87
8.2.	MÉTODO	91
8.3.	RESULTADOS	99
8.4.	DISCUSIÓN.....	114
8.5.	CONCLUSIONES	131
9.	RESERVAS FUSIONALES.....	133
9.1.	INTRODUCCIÓN	133
9.2.	MÉTODO	136
9.3.	RESULTADOS	140
9.4.	DISCUSIÓN.....	152
9.5.	CONCLUSIONES	158
10.	FLEXIBILIDAD DE VERGENCIAS	159
10.1.	INTRODUCCIÓN	159
10.2.	MÉTODO	160
10.3.	RESULTADOS	161
10.4.	DISCUSIÓN.....	162
10.5.	CONCLUSIONES.....	165
11.	ESTEREOAGUDEZA	167
11.1.	INTRODUCCIÓN	167
11.2.	MÉTODO	170
11.3.	RESULTADOS	178
11.4.	DISCUSIÓN.....	185
11.5.	CONCLUSIONES.....	189
12.	CONCLUSIONES Y PERSPECTIVAS FUTURAS.....	191
12.1.	CONCLUSIONES	191
12.2.	PROPUESTAS PARA INVESTIGACIONES FUTURAS.....	197
	REFERENCIAS	201
	ANEXOS.....	221
	ANEXO I. CONSENTIMIENTO INFORMADO.....	221
	ANEXO II. FICHAS DE REGISTRO DE LOS RESULTADOS DE LAS PRUEBAS	223
	ANEXO III. INFORME ENTREGADO A LOS PARTICIPANTES SOBRE SU ESTADO VISUAL	229
	ANEXO IV. MÉTODO DE BLAND Y ALTMAN	231
	ANEXO V. COMPARACIÓN DE MÉTODOS DE MEDIDA.....	237
	ANEXO VI. CONTRIBUCIONES CIENTÍFICAS DERIVADAS DEL DESARROLLO DEL DOCTORADO	249
	ANEXO VII. PROYECTOS DE INVESTIGACIÓN RELACIONADOS CON EL DESARROLLO DEL DOCTORADO	251

LISTA DE SÍMBOLOS, ABREVIATURAS Y SIGLAS

AA	Amplitud de acomodación
AC/A	(Cociente AC/A). Cambio en la convergencia acomodativa que ocurre al estimular o relajar la acomodación
ARN	Acomodación relativa negativa
ARP	Acomodación relativa positiva
AutoRx	Autorrefractómetro
AV	Agudeza visual
BE	Base externa
BI	Base interna
CA/A	(Cociente CA/A). Cambio en la acomodación que ocurre al estimular o relajar la convergencia
CCI	Coefficiente de correlación intraclase
CDC	Coefficiente de concordancia
CDR	Coefficiente de repetibilidad
CCF	Cilindros cruzados fusionados
cm	Centímetros
cpm	Ciclos por minuto
D	Dioptías
DE	Desviación estándar
DM	Diferencia media
<i>et al.</i>	Y colaboradores
FA	Flexibilidad acomodativa
FAB	Flexibilidad acomodativa binocular
FAM	Flexibilidad acomodativa monocular

FV	Flexibilidad de vergencias
m	Metros
MEM	Método de estimación monocular
min	Minutos
mseg	Milisegundos
PPC	Punto próximo de convergencia
r	Coefficiente de correlación de Pearson
RA	Respuesta acomodativa
s	Segundos
seg arc	Segundos de arco
VFN	Vergencia fusional negativa
VFP	Vergencia fusional positiva
VL	Visión lejana
VP	Visión próxima
vs.	<i>Versus</i> (respecto a, en comparación con)
Δ	Dioptrías prismáticas

GLOSARIO DE TÉRMINOS USADOS EN ESTUDIOS DE MEDICIÓN

La terminología usada en estudios de medición está llena de sinónimos, polisemias y términos confusos. Tal y como plantea Last¹ en su diccionario de epidemiología existen a veces dudas con respecto a los términos empleados para describir las propiedades de las mediciones: exactitud, precisión, validez, fiabilidad, repetibilidad y reproducibilidad. La exactitud y la precisión se utilizan a menudo de forma sinónima; la validez se define de distintos modos, y la fiabilidad, repetibilidad y reproducibilidad se usan con frecuencia de manera indistinta. Este glosario tiene por objeto concretar el significado con el que se utilizan los diferentes términos a lo largo de esta tesis^{2,3}.

Concordancia: (sinónimo de Acuerdo). Cuando existen diferentes métodos de medida para una misma variable, esta propiedad establece hasta qué punto los resultados obtenidos con distintos métodos son equivalentes. Es decir, evalúa variabilidad de los resultados entre dos pruebas que pretenden medir la misma variable.

Consistencia: Estrecha uniformidad entre los hallazgos que procedan de diferentes muestras o poblaciones, obtenidos en momentos distintos, en distintas circunstancias o en estudios llevados a cabo por diversos métodos o investigadores. La consistencia puede examinarse para estudiar la modificación de los efectos.

Error: Resultado falso o equivocado obtenido en un estudio o experimento. Pueden darse diversas clases de errores. Error aleatorio: Es aquella porción de variación de una medida que no tiene relación aparente con ninguna otra medida o variable, y que se considera debido al azar. Error sistemático: porción de variación de una medida que a menudo tiene un origen reconocible, por ejemplo, un instrumento defectuoso o un patrón erróneo.

Fiabilidad: Se refiere al grado en que pueden reproducirse los resultados obtenidos por un procedimiento de medición cuando se aplica a la misma persona en más de una ocasión, pero en idénticas condiciones. Es un concepto muy relacionado con la variabilidad aleatoria. Cuanta menor variabilidad se obtenga al repetir una medición, más fiable será. La falta de fiabilidad puede surgir de divergencias entre los observadores o de los instrumentos de medición o por inestabilidad del atributo objeto de medición y en la práctica, todas las fuentes de variación pueden actuar de forma simultánea. Todas las observaciones se hallan sujetas a variación. Es de esperar que existan discrepancias entre observaciones repetidas efectuadas por el

mismo observador, así como entre diferentes observadores; dichas discrepancias pueden minimizarse, pero probablemente nunca pueden eliminarse del todo. Un modo de descubrir si una medición es o no fiable consiste en repetirla o reproducirla, de ahí que fiabilidad, repetibilidad y reproducibilidad tiendan a utilizarse indistintamente. Se evalúa la fiabilidad intra-examinador cuando se establece el grado de consistencia al efectuar la medición de un observador consigo mismo; mientras que la fiabilidad inter-examinador se refiere a la consistencia entre varios observadores cuando efectúan la medición en un mismo individuo. Para asegurar la independencia de las medidas de los distintos examinadores, un observador no deben conocer los resultados proporcionados por los demás (observadores ciegos). Corresponde a la Precisión de la medida.

Medición: Proceso mediante el cual se cuantifica una magnitud por comparación con una escala preestablecida.

Repetibilidad: (sinónimo de Reproducibilidad): Capacidad del método de medición para producir el mismo resultado (o uno estrechamente semejante) cada vez que se aplica el procedimiento en condiciones idénticas. Por paralelismo con la fiabilidad, se puede decir que un determinado procedimiento presenta una buena repetibilidad intra-examinador si tiene a producir el mismo resultado cuando se aplica por el mismo examinador en sucesivas ocasiones a un mismo paciente; y que presenta una buena repetibilidad inter-examinador cuando tiende a ofrecer el mismo resultado cuando se aplica por varios observadores en un mismo individuo.

Sesgo: Desviación de los resultados o de las deducciones respecto a la verdad.

Validez: Propiedad de un método de medida que indica en qué grado el procedimiento mide la magnitud que se desea cuantificar. Dicho de otro modo, expresa el grado en que el valor que se obtiene con el proceso de medición se corresponde con el verdadero valor de la variable en un sujeto, es decir, que no existe un sesgo (error sistemático) en el proceso de medición. Las fuentes de error sistemático son las mismas que influyen sobre la fiabilidad. No existen instrumentos perfectamente válidos, de modo que se trata siempre de una cuestión de grado. Se suele estimar por comparación con una prueba patrón de referencia (*Gold Standard*) que deberá ser una prueba repetible y ampliamente aceptada como buena para la medida del fenómeno de interés. Corresponde a la Exactitud de la medida.

Variable: Todo atributo, fenómeno o hecho que pueda tener diferentes valores.

LISTA DE TABLAS

Tabla 1. Estado actual del tema en el ámbito de los estudios de repetibilidad y concordancia de las pruebas que evalúan las habilidades visuales binoculares	6
Tabla 2. Repetibilidad intra-examinador de las pruebas de visión binocular.....	10
Tabla 3. Prueba recomendada para la evaluación de cada una de las habilidades visuales binoculares estudiadas	11
Tabla 4. Concordancia entre las pruebas que miden la misma habilidad visual	13
Tabla 5. Número de sujetos necesarios para la estimación de la concordancia intraobservador (Tabla H de la sección de anexos del libro de Argimon y Jiménez ³).	27
Tabla 6. Características principales de la muestra. Género, medias \pm desviación estándar (DE) de la agudeza visual decimal equivalente de Snellen (AV) y de diferentes parámetros relacionados con las ametropías.	31
Tabla 7. Pruebas estudiadas que se clasifican según la cualidad visual que miden ...	33
Tabla 8. Repetibilidad en la medida de la amplitud de acomodación.....	44
Tabla 9. Concordancia entre pruebas que miden la amplitud de acomodación (AA)...	46
Tabla 10. Repetibilidad en la medida de flexibilidad acomodativa	60
Tabla 11. Repetibilidad en la medida de la respuesta acomodativa (RA).	75
Tabla 12. Concordancia entre pruebas que miden la respuesta acomodativa (RA). ...	77
Tabla 13. Repetibilidad en la medida de acomodación relativa.	85
Tabla 14. Pruebas de medida de la desviación ocular estudiadas.....	91
Tabla 15. Repetibilidad en la medida de desviación ocular horizontal en visión lejana y en visión próxima	101
Tabla 16. Repetibilidad en la medida de desviación ocular vertical en cerca	105
Tabla 17. Concordancia entre pruebas que miden la desviación ocular horizontal ...	108
Tabla 18. Concordancia entre pruebas que miden la desviación ocular vertical.....	113
Tabla 19. Resumen de estudios de repetibilidad en la medida de forias	115

Tabla 20. Resumen de estudios que comparan las pruebas de medida de forias.....	118
Tabla 21. Repetibilidad en la medida del PPC	141
Tabla 22. Repetibilidad de la medida de las reservas fusionales horizontales con barra de prismas y con los diasporámetros del foróptero.	143
Tabla 23. Concordancia de la medida de vergencias fusionales con diasporámetro y con barra de prismas	149
Tabla 24. Repetibilidad en la medida de flexibilidad de vergencias	161
Tabla 25. Descripción de los tests de estereoagudeza	171
Tabla 26. Repetibilidad en la medida de la estereoagudeza en visión próxima	178
Tabla 27. Repetibilidad en la medida de estereoagudeza en visión próxima por tramos	180
Tabla 28. Concordancia entre pruebas que miden la estereoagudeza en visión próxima	182
Tabla 29. Concordancia del test TNO con los demás estereotests por tramos	184
Tabla 30. Resultados de la medición de cierta cualidad utilizando dos métodos A y B (ejemplo).	234
Tabla 31. Valoración de la concordancia según los valores del coeficiente de correlación intraclase (CCI).....	242

LISTA DE FIGURAS

Figura 1. Extracción de bolas numeradas para establecer aleatoriamente el orden de las pruebas.....	36
Figura 2. Medida de la amplitud de acomodación por acercamiento (izquierda) y por alejamiento (derecha)	42
Figura 3. Medida de la amplitud de acomodación con el método de las lentes negativas	43
Figura 4. Diferencia media (DM), límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) y coeficiente de repetibilidad (CDR) para la medida de amplitud de acomodación (AA). En la AA por acercamiento la distribución de las diferencias no fue normal y se tomó como coeficiente de repetibilidad el percentil 95 del valor absoluto de las diferencias (CDR*).....	45
Figura 5. Gráficas de Bland y Altman para el estudio de la concordancia entre métodos que miden la amplitud de acomodación (AA). La línea continua representa la diferencia media (DM) entre las medidas obtenidas con los dos métodos. Las líneas discontinuas indican los límites superior (S) e inferior (I) del intervalo de acuerdo al 95% ($DM \pm 1,96 DE$). Cuando la distribución de las diferencias no fue normal, se tomó como coeficiente de concordancia el percentil 95 del valor absoluto de las diferencias (CDC*).....	47
Figura 6. Medida de la flexibilidad acomodativa con flipper de $\pm 2,00 D$	58
Figura 7. Diferencia media (DM), límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) y coeficiente de repetibilidad (CDR) para la medida de la flexibilidad acomodativa binocular (FAB) y monocular (FAM). En la FAB la distribución de las diferencias no fue normal y se tomó como coeficiente de repetibilidad el percentil 95 del valor absoluto de las diferencias (CDR*).....	59
Figura 8. Medida de la respuesta acomodativa con retinoscopía Nott.....	69
Figura 9. Medida de la respuesta acomodativa con retinoscopía MEM	70
Figura 10. Medida de la respuesta acomodativa con el método de los cilindros cruzados fusionados. (Al tomar la fotografía se dio la vuelta a la tarjeta para que mostrar el estímulo utilizado).	72
Figura 11. Medida de la respuesta acomodativa con autorrefractómetro de campo abierto Shin-Nippon SRW-5000 con estímulo de fijación en visión próxima.	73

Figura 12. Diferencia media (DM), límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) y coeficiente de repetibilidad (CDR) para la medida de la respuesta acomodativa (RA). En la RA con el método de los cilindros cruzados fusionados (CCF) y con autorrefractómetro en cerca la distribución de las diferencias no fue normal y se tomó como coeficiente de repetibilidad el percentil 95 del valor absoluto de las diferencias (CDR*).	74
Figura 13. Gráficas de Bland y Altman para el estudio de la concordancia entre métodos que miden la respuesta acomodativa. La línea continua representa la diferencia media (DM) entre las medidas obtenidas con los distintos métodos. Las líneas discontinuas indican los límites superior (S) e inferior (I) del intervalo de acuerdo al 95% ($DM \pm 1,96 \cdot DE$). Cuando la distribución de las diferencias no fue normal, se tomó como coeficiente de concordancia el percentil 95 del valor absoluto de las diferencias (CDC*).	76
Figura 14. Medida de la acomodación relativa negativa y positiva (ARN y ARP)	84
Figura 15. Diferencia media (DM) y coeficiente de repetibilidad (CDR) para la medida de la acomodación relativa negativa (ARN) y positiva (ARP). En ambos casos la distribución de las diferencias no fue normal y se tomó como coeficiente de repetibilidad el percentil 95 del valor absoluto de las diferencias (CDR*) y se anotaron los límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I).	85
Figura 16. Medida de la desviación ocular con la técnica del prisma cover test en visión próxima (izquierda) y en visión lejana (derecha).	93
Figura 17. Punto de para la medida de forias horizontales con la técnica de von Graefe utilizando el foróptero.....	94
Figura 18. Medida de forias con varilla de Maddox en visión próxima utilizando el foróptero.....	96
Figura 19. Medida de forias en visión próxima con el método de Thorington modificado.	97
Figura 20. Tarjeta Bernell Muscle Imbalance Measure (MIM) utilizada para medir la foria en visión próxima.	99
Figura 21. Tendencia de valores medios de las distintas técnicas de medida de desviación horizontal.....	101
Figura 22. Diferencia media (DM), límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) y coeficiente de repetibilidad (CDR) para la medida de la desviación ocular horizontal en visión lejana.	103
Figura 23. Diferencia media (DM), límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) y coeficiente de repetibilidad (CDR) para la desviación ocular horizontal en visión próxima.....	104

- Figura 24. Diferencia media (DM), límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) y coeficiente de repetibilidad (CDR) para la desviación ocular vertical en visión próxima. En los tres métodos la distribución de las diferencias no fue normal y se tomó como coeficiente de repetibilidad el percentil 95 del valor absoluto de las diferencias (CDR*)..... 106
- Figura 25. Gráficas de Bland y Altman para el estudio de la concordancia entre métodos que miden la foria horizontal en lejos (6 m). La línea continua representa la diferencia media (DM) entre las medidas obtenidas con los distintos métodos. Las líneas discontinuas indican los límites superior (S) e inferior (I) del intervalo de acuerdo al 95% ($DM \pm 1,96 \cdot DE$). 109
- Figura 26. Gráficas de Bland y Altman para el estudio de la concordancia entre métodos que miden la foria horizontal en cerca (40 cm). La línea continua representa la diferencia media (DM) entre los distintos métodos. Las líneas discontinuas indican los límites superior (S) e inferior (I) del intervalo de acuerdo al 95% ($DM \pm 1,96 \cdot DE$). 110
- Figura 27. Gráficas de Bland y Altman para el estudio de la concordancia entre métodos que miden la foria vertical en visión próxima (40 cm). La línea continua representa la diferencia media (DM) entre métodos. Las líneas discontinuas indican los límites superior (S) e inferior (I) del intervalo de acuerdo al 95% ($DM \pm 1,96 DE$). En todos los casos la distribución de las diferencias no fue normal y se tomó como coeficiente de concordancia el percentil 95 del valor absoluto de las diferencias (CDC*). 112
- Figura 28. Medida de las vergencias fusionales horizontales con diasporámetro 138
- Figura 29. Medida de las reservas fusionales horizontales con barra de prismas 139
- Figura 30. Medida del punto próximo de convergencia (PPC). 140
- Figura 31. Diferencia media (DM), límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) y coeficiente de repetibilidad (CDR) para la medida del punto próximo de convergencia (PPC). Tanto en la rotura (Ro) como en el recobro (Re) la distribución de las diferencias no fue normal y se tomó como coeficiente de repetibilidad el percentil 95 del valor absoluto de las diferencias (CDR*)..... 141
- Figura 32. Diferencia media (DM), límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) y coeficiente de repetibilidad (CDR) para la medida de las reservas fusionales (VFN y VFP) con barra de prismas en lejos. Cuando la distribución de las diferencias no fue normal y se tomó como coeficiente de repetibilidad el percentil 95 del valor absoluto de las diferencias (CDR*)..... 144
- Figura 33. Diferencia media (DM), límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) y coeficiente de repetibilidad (CDR) para la medida de las reservas fusionales (VFN y VFP) con diasporámetro en lejos.

Quando la distribución de las diferencias no fue normal y se tomó como coeficiente de repetibilidad el percentil 95 del valor absoluto de las diferencias (CDR*)..... 145

Figura 34. Diferencia media (DM), límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) y coeficiente de repetibilidad (CDR) para la medida de las reservas fusionales (VFN y VFP) con barra de prismas en cerca. Cuando la distribución de las diferencias no fue normal y se tomó como coeficiente de repetibilidad el percentil 95 del valor absoluto de las diferencias (CDR*)..... 146

Figura 35. Diferencia media (DM), límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) y coeficiente de repetibilidad (CDR) para la medida de las reservas fusionales (VFN y VFP) con diasporámetro en cerca. Cuando la distribución de las diferencias no fue normal y se tomó como coeficiente de repetibilidad el percentil 95 del valor absoluto de las diferencias (CDR*)..... 147

Figura 36. Gráficas de Bland y Altman para el estudio de la concordancia entre métodos que miden las reservas fusionales horizontales en lejos (6 m). La línea continua representa la diferencia media (DM) entre las medidas obtenidas con diasporámetro y con barra de prismas. Las líneas discontinuas indican los límites superior (S) e inferior (I) del intervalo de acuerdo al 95% ($DM \pm 1,96 \cdot DE$). Cuando la distribución de las diferencias no fue normal, se tomó como coeficiente de concordancia el percentil 95 del valor absoluto de las diferencias (CDC*)..... 150

Figura 37. Gráficas de Bland y Altman para el estudio de la concordancia entre métodos que miden las reservas fusionales horizontales en cerca (40 cm). La línea continua representa la diferencia media (DM) entre las medidas obtenidas con diasporámetro y con barra de prismas. Las líneas discontinuas indican los límites superior (S) e inferior (I) del intervalo de acuerdo al 95% ($DM \pm 1,96 \cdot DE$). Cuando la distribución de las diferencias no fue normal, se tomó como coeficiente de concordancia el percentil 95 del valor absoluto de las diferencias (CDC*)..... 151

Figura 38. Correlación entre la medida del PPC y la VFP en VP medida con barra de prismas (izquierda) y con diasporámetro (derecha)..... 152

Figura 39. Diferencia media (DM), límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) y coeficiente de repetibilidad (CDR) para la medida de flexibilidad de vergencias (FV) con flippers de 3Δ BI / 12Δ BE. La distribución de las diferencias no fue normal y se tomó como coeficiente de repetibilidad el percentil 95 del valor absoluto de las diferencias (CDR*)..... 162

Figura 40. Correlación entre las medidas de flexibilidad de vergencias inicial y final. 164

Figura 41. Estereoagudeza con el test de Frisby 173

Figura 42. Estereoagudeza con el test TNO 174

Figura 43. Estereoagudeza con el test de Randot 176

Figura 44. Estereoagudeza con el test de Titmus 177

-
- Figura 45. Diferencia media (DM) y límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) para los diferentes tests de medida de la estereoagudeza en cerca. En todos los casos la distribución de las diferencias no fue normal y se tomó como coeficiente de repetibilidad el percentil 95 del valor absoluto de las diferencias (CDR*). 179
- Figura 46. Gráficas de Bland y Altman para el estudio de la concordancia entre métodos que miden la estereoagudeza en visión próxima. La línea continua representa la diferencia media (DM) entre las medidas obtenidas con los distintos métodos. Las líneas discontinuas indican los límites superior (S) e inferior (I) del intervalo de acuerdo al 95% ($DM \pm 1,96 DE$). En todos los casos la distribución de las diferencias no fue normal y se tomó como coeficiente de concordancia el percentil 95 del valor absoluto de las diferencias (CDC*). 183
- Figura 47. Gráfica de Bland y Altman correspondiente al estudio de concordancia entre los métodos de medida A y B (ejemplo). 235
- Figura 48. Gráfica de correlación entre los métodos de medida A y B (ejemplo) 238

PARTE I:

Aspectos Generales

Introducción

La medida clínica de calidad es la condición básica tanto de la investigación clínica de excelencia, como de la correcta atención a los pacientes, pues sin ella los cambios clínicos, los éxitos, o los fracasos de un tratamiento no pueden ser correctamente detectados. La atención clínica en optometría se realiza mediante la entrevista, la observación y las pruebas clínicas. Las características más importantes a tener en cuenta al aplicar las pruebas son su repetibilidad, su validez y su capacidad discriminativa. La repetibilidad parece ser la cualidad más importante de una prueba, porque influye en las demás. Las pruebas deben tener alta repetibilidad para que puedan ser válidas y para que tengan buena capacidad discriminativa.

Uno de los problemas que se presenta al clínico que quiere elegir una determinada prueba, entre las diversas que existen con la misma finalidad, es el desconocimiento de la repetibilidad de dichas pruebas. Interesa conocer si las dos pruebas son comparables o si existen diferencias que permitan afirmar la superioridad de una de ellas sobre la otra en términos de repetibilidad.

Tal y como expresa Bullimore⁴ en su editorial no hay duda de que los estudios de repetibilidad aportan información importante tanto para los clínicos como para los investigadores. Esos resultados nos dicen que algunas pruebas tienen elevada repetibilidad y, en cambio, otras presentan una variabilidad preocupante. Con estos estudios, el clínico conocerá las limitaciones de sus medidas y estará más orientado a la hora de adquirir futuros instrumentos. Por su parte, los investigadores elegirán las pruebas más adecuadas al planificar sus investigaciones. Los clínicos somos personas de costumbres y solemos elegir unas determinadas pruebas que aplicamos exactamente igual a nuestros pacientes durante años. Sin embargo, hay ocasiones

donde los resultados de un estudio de repetibilidad nos permiten darnos cuenta de que otras técnicas disponibles presentan mejor desempeño y que nuestro modo tradicional de hacer las cosas no siempre es el mejor.

Existen, al menos, cuatro áreas en las que la evaluación del cambio en los resultados de las pruebas puede ser importante en la toma de decisiones clínicas:

- (1) La detección de enfermedades
- (2) El control de la remisión o de la progresión de la enfermedad
- (3) El control del efecto de los tratamientos
- (4) La evaluación de las diferencias en el desempeño bajo diferentes condiciones de prueba⁵.

A pesar de la obvia importancia de valorar el cambio en los resultados de cada prueba, hay poca normativa sobre la significación del cambio y cuando existe, raramente se utiliza. Consecuentemente, los profesionales juzgan la importancia clínica del cambio experimentado por un paciente en determinada prueba basándose en su propia intuición.

Revisando las publicaciones del área biosanitaria se observa que cada vez existen más estudios sobre la repetibilidad de los tests. Sin embargo, en el ámbito de las anomalías binoculares (acomodación, convergencia) no existen muchos trabajos, no hay siempre consenso entre sus conclusiones y sólo recientemente están analizándose los datos con las técnicas estadísticas adecuadas. Fruto de esta situación, que será analizada con más detalle en el capítulo "FUNDAMENTACIÓN TEÓRICA" y en el sub-apartado "INTRODUCCIÓN" de cada uno de los experimentos,

surgió la idea de plantear este estudio cuyos objetivos se plantean en el siguiente apartado.

Objetivos de esta investigación

El presente trabajo se enmarca dentro de la línea de investigación sobre evaluación de pruebas optométricas desarrollada por el equipo de investigación de Optometría Infantil, Visión Binocular y Ortóptica del Departamento Óptica II (Optometría y Visión) de la UCM, dirigido por el profesor Barra. El **principal objetivo** de esta investigación fue estudiar la repetibilidad de las pruebas de visión binocular más habitualmente utilizadas dentro del examen optométrico en una población de adultos jóvenes con altas demandas visuales y visión binocular normal.

En primer lugar, se quiso averiguar la **repetibilidad** de las pruebas sometidas a estudio. En concreto, el estudio se centró en la repetibilidad intra-examinador que es la repetibilidad de una prueba para un único examinador utilizando la misma metodología en varias ocasiones diferentes. Para ello, de cada prueba se tomaron medidas en dos ocasiones diferentes separadas temporalmente y se compararon entre sí. Para que la repetibilidad de una prueba sea alta la diferencia entre las dos medidas debe ser lo menor posible en comparación con el rango típico de resultados de la variable evaluada mediante dicha prueba. Si las medidas hubieran sido tomadas por varios examinadores se estaría introduciendo otra fuente de variabilidad y cabría esperar un aumento del error de la medida. Probablemente este aumento del error no sería importante en aquellas pruebas en las que la intervención del examinador en la medida es pequeña (pruebas subjetivas) y sería más elevado en las pruebas objetivas⁵.

En segundo lugar, puesto que se han descrito variadas pruebas optométricas para evaluar la misma habilidad visual, se quiso comparar entre sí los resultados ofrecidos por cada una de las pruebas destinadas a medir el mismo parámetro. De este modo se pudo establecer el nivel de **concordancia** entre los resultados de los distintos procedimientos orientados a caracterizar determinada habilidad y establecer si se pueden o no considerar pruebas intercambiables o por el contrario parecen medir cualidades distintas.

Aunque a lo largo de los años se han propuesto numerosos métodos clínicos y experimentales de evaluación del estado binocular, este trabajo se centró en analizar las pruebas más ampliamente utilizadas en la práctica clínica habitual. En este trabajo se propusieron los siguientes **objetivos concretos**:

- (1) Estandarizar las condiciones de contorno y la metodología a aplicar al administrar las pruebas estudiadas.
- (2) Determinar la repetibilidad intra-examinador de los tests.
- (3) Comparar la repetibilidad de los distintos tests.
- (4) Determinar si hubo alguna diferencia sistemática (o sesgo relativo) entre los distintos métodos que miden la misma cualidad visual.
- (5) Establecer el grado de concordancia entre parejas de tests que miden la misma variable.

La **hipótesis general** del estudio es que las pruebas tienen distintos valores de repetibilidad y que algunas de ellas tienen mejores valores de esta característica. Se espera que los métodos objetivos muestren mejor repetibilidad que los subjetivos. En cuanto a la concordancia, la hipótesis general es que dentro de cada grupo de pruebas

que evalúan la misma variable, unas parejas de métodos tienen mejor concordancia entre sí que otras parejas. Además, si la repetibilidad de uno de los métodos de la pareja es baja, el grado de concordancia entre ambos será probablemente bajo también y si ambos métodos de la pareja tienen escasa repetibilidad, la situación será aun peor. Ambas hipótesis se han puesto a prueba en esta investigación, particularmente para cada conjunto de métodos que miden la misma variable: amplitud de acomodación, valor de la heteroforia, etc.

Aportaciones novedosas

El interés de esta tesis surge, en parte, debido a la escasez de investigación existente sobre la repetibilidad de muchas de las pruebas de estudio de la visión binocular aquí estudiadas. La Tabla 1 resume el trabajo de revisión sobre el estado actual del tema de la tesis doctoral y posteriormente en los apartados “Introducción” y “Discusión” de cada experimento se discutirán con detalle los resultados concretos de las investigaciones aludidas. En la Tabla 1 se puede ver que el presente trabajo es el primero que estudia la repetibilidad de varias de las pruebas analizadas, por ejemplo, en la evaluación de la acomodación relativa negativa y positiva, en la medida del retraso acomodativo con los cilindros cruzados fusionados y en la medida de la estereoagudeza con el test TNO y con el test de Titmus. En otros casos, es el primer estudio en el que se han utilizado los protocolos estadísticos más apoyados en la actualidad para establecer la repetibilidad de una prueba clínica y el grado de concordancia entre distintas pruebas que evalúan la misma habilidad visual, tal es el caso de la amplitud de acomodación por el método de las lentes negativas o por el de alejamiento y el de la flexibilidad acomodativa con flippers. Por último, hay algunas pruebas en las que la repetibilidad sólo ha sido correctamente establecida en un rango de edad diferente al de la muestra del presente estudio. Así, la medida de la

flexibilidad de vergencias⁶ se ha estudiado en presbítas, mientras que la medida del punto próximo de convergencia (PPC)⁷, de la amplitud de acomodación por acercamiento⁷ y de la flexibilidad acomodativa⁸⁻¹⁰ sólo se ha estudiado correctamente en niños.

Tabla 1. Estado actual del tema en el ámbito de los estudios de repetibilidad y concordancia de las pruebas que evalúan las habilidades visuales binoculares

HABILIDAD ESTUDIADA	ESTADO ACTUAL DEL TEMA
AMPLITUD DE ACOMODACIÓN	<ul style="list-style-type: none"> - No hay trabajos previos que utilicen los métodos estadísticos recomendados para estudiar la repetibilidad del método de las lentes negativas. - La técnica de acercamiento sólo se ha analizado de forma correcta en niños - No hay estudios de concordancia que apliquen el análisis estadístico adecuado.
FLEXIBILIDAD ACOMODATIVA	<ul style="list-style-type: none"> - No hay trabajos previos que utilicen la metodología recomendada
RETRASO ACOMODATIVO	<ul style="list-style-type: none"> - Ha sido un aspecto algo más estudiado y se han encontrado trabajos previos de los distintos métodos excepto de la prueba de los cilindros cruzados fusionados (CCF)
ACOMODACIÓN RELATIVA	<ul style="list-style-type: none"> - No hay ningún trabajo que estudie la repetibilidad de estas pruebas
DESVIACIÓN OCULAR	<ul style="list-style-type: none"> - Es posiblemente una de las habilidades visuales más estudiada en los últimos 50 años, sin embargo se ha observado una importante falta de homogeneidad metodológica que dificulta notoriamente la comparación de resultados de las distintas investigaciones.
RESERVAS FUSIONALES HORIZONTALES	<ul style="list-style-type: none"> - Sólo se han aplicado los métodos estadísticos recomendados en la medida del punto de rotura base externa con diasporámetro y con la prueba del PPC, pero no se ha estudiado la medida base interna, ni el emborronamiento o el recobro base externa y además, se llevó a cabo con niños. - No se ha estudiado la medida de reservas fusionales con barra de prismas. - No hay estudios de concordancia
FLEXIBILIDAD DE VERGENCIAS	<ul style="list-style-type: none"> - Sólo se ha realizado un análisis estadístico correcto en presbítas.
ESTEREOAGUDEZA	<ul style="list-style-type: none"> - No hay estudios de repetibilidad intra-examinador para el test TNO ni para el test de Titmus. - No hay estudios de concordancia que apliquen el análisis estadístico adecuado.

Tras revisar las publicaciones relacionadas con el propósito de este trabajo se ha detectado una falta de coincidencia sobre la forma de abordar el análisis estadístico en este tipo de estudios; por lo que esta tesis aporta el hecho de estudiar todas las pruebas bajo el mismo patrón.

Está claramente justificada la conveniencia de un estudio cuyos resultados aporten información que permita establecer la prueba más adecuada para su uso clínico desde el punto de vista de su repetibilidad. Las características más importantes que se pueden encontrar en los artículos de investigación que describen la utilidad de una prueba clínica son su validez, su capacidad de discriminación y su repetibilidad intra-examinador e inter-examinador¹¹. La prueba ideal debe tener buena repetibilidad, una buena validez, una buena capacidad discriminativa y ser fácil de administrar. De estos aspectos, la presente investigación se centró en tratar de determinar cuales de las pruebas utilizadas para evaluar la visión binocular presentan una buena repetibilidad intra-examinador.

Las principales aportaciones de la presente investigación al ámbito de la práctica clínica, han sido las siguientes:

- **Descripción detallada de cada procedimiento**

En el apartado “método” de los distintos capítulos se ha descrito con detalle el protocolo seguido al administrar cada una de las pruebas estudiadas. Para su preparación se han utilizado como fuentes de información los principales manuales optométricos¹¹⁻¹⁶ así como las instrucciones que acompañan al instrumento o test cuando procedía. Por lo que de esta tesis se puede desprender un pequeño manual de instrucciones en el que se estandarizan las pruebas clínicas de visión binocular más habituales.

- **Repetibilidad intra-examinador.**

La Tabla 2 refleja la diferencia media (DM) y el coeficiente de repetibilidad (CDR) (ver apartado 3.3. “*Tratamiento estadístico de los datos*”) que corresponde al mínimo cambio clínicamente significativo para cada una de las pruebas estudiadas para una población clínicamente normal de adultos jóvenes (18 a 32 años). Los CDRs además de permitir comparar las distintas pruebas entre sí, representan el límite inferior de las variaciones potenciales intra-examinador.

Desgraciadamente cuando un clínico trata y controla determinada alteración para poder tener la confianza de que el cambio que se ha producido en la cualidad medida en un paciente es un cambio real y no una variación natural y esperable entre distintas sesiones de medida, será necesario que utilice las diferencias extremas limitadas por los límites de repetibilidad. Cuanto más grandes sean las diferencias test-retest más difícil será tomar las decisiones clínicas para fijar un diagnóstico y un tratamiento.

Teniendo en cuenta los amplios CDR de varias de las pruebas evaluadas, cabe reflexionar sobre si se deben considerar los datos medidos a un paciente en un examen clínico desde la perspectiva de las mayores diferencias respecto a los datos medidos que pueden existir debidos a errores naturales del proceso de medida. Respecto a esto hay que tener en cuenta que la mayoría de los pacientes tienen un comportamiento cercano a lo típico pero, por supuesto, hay unos pocos pacientes problemáticos que no cumplen esta expectativa. Sugiero analizar los datos clínicos de la mayoría de los pacientes utilizando las diferencias medias (DM) encontradas para cada prueba y preguntarse si se alteraría el diagnóstico si las medidas observadas cambiaran según marcan los límites de acuerdo al 95% (CDR).

Será importante considerar las diferencias extremas cuando haya inconsistencias en el examen del paciente; por ejemplo, cuando unos datos clínicos apoyen el diagnóstico de determinada alteración, y sin embargo el paciente no presente síntomas o éstos sean muy leves en comparación de los esperados para ese diagnóstico. También se recomienda considerar las diferencias extremas cuando un paciente presente los síntomas característicos de determinada alteración, pero los resultados de las pruebas clínicas aplicadas se encuentren dentro de los rangos de normalidad.

Se debe tener en cuenta que los valores esperados de CDR basados en las diferencias entre sesiones para una población normal pueden llevar a concluir que para un sujeto con determinada alteración visual se ha producido un cambio significativo cuando puede que no sea así. Esto se debe a que cabe esperar que sujetos con una enfermedad estable presenten variaciones considerablemente mayores que los sujetos normales. Es decir, la estimación del error de medida establecido para una población clínicamente normal puede que no sea igual al que se extraería en individuos con alteración. Sin embargo, la aplicación de una tabla estándar de valores esperados de error para una población normal como la que se propone aquí (ver Tabla 2) es preferible a no tener en cuenta ningún criterio sobre variaciones debidas al error de medida cuando se toman decisiones sobre la significación del cambio en las respuestas de un paciente.

Tabla 2. Repetibilidad intra-examinador de las pruebas de visión binocular

HABILIDAD ESTUDIADA	DM±CDR	
AMPLITUD DE ACOMODACIÓN <ul style="list-style-type: none"> Acercamiento Alejamiento Lentes negativas 	-0,61±4,76 D -0,81±4,00 D -0,08±2,52 D	
FLEXIBILIDAD ACOMODATIVA <ul style="list-style-type: none"> Binocular con flipper Monocular con flipper 	0,85±6,50 cpm 1,16±5,43 cpm	
RESPUESTA ACOMODATIVA <ul style="list-style-type: none"> Retinoscopia Nott Retinoscopia MEM Cilindros cruzados fusionados Autorrefractómetro en cerca 	-0,10±0,66 D -0,23±0,98 D -0,05±0,75 D -0,12±1,00 D	
ACOMODACIÓN RELATIVA <ul style="list-style-type: none"> ARN con foróptero ARP con foróptero 	0,03±0,75 D -0,38±2,25 D	
DESVIACIÓN OCULAR HORIZONTAL <ul style="list-style-type: none"> Cover test (neutralización) Cover test (media neutr.-inversión) Técnica de von Graefe Varilla de Maddox Método de Thorington modificado 	LEJOS 0,10±1,17 Δ 0,12±2,84 Δ -0,43±3,98 Δ 0,53±4,53 Δ ---	CERCA -0,21±3,18 Δ -0,43±4,97 Δ -0,70±9,26 Δ -0,20±8,27 Δ -0,60±6,52 Δ
DESVIACIÓN OCULAR VERTICAL <ul style="list-style-type: none"> Técnica de von Graefe Varilla de Maddox Método de Thorington modificado 	CERCA 0,08±1,00 Δ -0,03±0,70 Δ 0,00±0,50 Δ	
RESERVAS FUSIONALES (LEJOS) <ul style="list-style-type: none"> Con barra de prismas <ul style="list-style-type: none"> VFN: E / Ro / Re VFP: E / Ro / Re Con diasporámetro <ul style="list-style-type: none"> VFN: E / Ro / Re VFP: E / Ro / Re RESERVAS FUSIONALES (CERCA) <ul style="list-style-type: none"> Con barra de prismas <ul style="list-style-type: none"> VFN: E / Ro / Re VFP: E / Ro / Re Con diasporámetro <ul style="list-style-type: none"> VFN: E / Ro / Re VFP: E / Ro / Re PPC: Ro / Re 	LEJOS --- / 0,20±4,00 / 0,36±4,00 Δ 1,17±9,73 / 3,85±13,93 / 2,44±8,30 Δ --- / 0,05±4,82 / 0,47±3,00 Δ -0,54±14,00 / 2,97±11,29 / 2,63±12,00 Δ CERCA 1,54±6,60 / -0,08±8,05 / -0,21±6,00 Δ 2,16±10,21 / 1,30±15,00 / 2,22±11,50 Δ 0,39±6,43 / -0,05±7,00 / -0,44±7,49 Δ 1,17±8,83 / 2,08±7,74 / 4,02±16,40 Δ 0,13±3,20 / 0,03±2,77 cm	
FLEXIBILIDAD DE VERGENCIAS <ul style="list-style-type: none"> Con flipper de 3Δ BI / 12Δ BE 	1,96±5,29 cpm	
ESTEREOAGUDEZA <ul style="list-style-type: none"> Test de Frisby Test TNO Test de Randot Test de Titmus 	-2,30±35 seg arc -3,52±60 seg arc -1,23±25 seg arc 2,38±40 seg arc	
Clave: DM = diferencia media (Final – Inicial). CDR = coeficiente de repetibilidad. D = dioptría. Δ = dioptría prismática. cpm = ciclos por minuto. MEM = método de estimación monocular. ARN/ARP = acomodación relativa negativa/positiva. VFN/VFP = vergencia fusional negativa/positiva. E = emborronamiento. Ro = rotura. Re = recobro		

Este estudio sólo considera la variación entre medidas tomadas en una visita inicial comparadas con una segunda visita. En la práctica clínica, durante el seguimiento de un paciente se pueden tener datos de más de dos sesiones y es importante darse cuenta de que siendo estrictos, los resultados del presente trabajo no serían aplicables a esta situación. Sin embargo, en ausencia de otra normativa más apropiada, los resultados extraídos de este estudio pueden ser una buena guía, que siempre será mejor que no tener en cuenta ningún dato de repetibilidad de las pruebas.

- **Pruebas recomendadas.**

En la Tabla 3 se recoge la prueba recomendada para la evaluación de cada una de las habilidades estudiadas, basándonos en su mayor repetibilidad intra-examinador.

Tabla 3. Prueba recomendada para la evaluación de cada una de las habilidades visuales binoculares estudiadas

HABILIDAD ESTUDIADA	PRUEBA RECOMENDADA
AMPLITUD DE ACOMODACIÓN	Lentes negativas
FLEXIBILIDAD ACOMODATIVA	Flipper de $\pm 2,00$ D
RETRASO ACOMODATIVO	Retinoscopía Nott
ACOMODACIÓN RELATIVA	ARN/ARP con foróptero
DESVIACIÓN OCULAR HORIZONTAL	Cover test (neutralización) Método de Thorington modificado
DESVIACIÓN OCULAR VERTICAL (VP)	Método de Thorington modificado
RESERVAS FUSIONALES HORIZONTALES	Diasporámetro
FLEXIBILIDAD DE VERGENCIAS	Flipper de 3Δ BI / 12Δ BE
ESTEREOAGUDEZA	Test de Randot

- **Concordancia entre pruebas.**

La Tabla 4 refleja la diferencia media (DM) y el coeficiente de concordancia (CDC) (ver apartado 3.3. “*Tratamiento estadístico de los datos*”) encontrado al comparar las diferentes pruebas que evalúan la misma habilidad visual para una población clínicamente normal de adultos jóvenes (18 a 32 años). La amplitud de los CDCs permite establecer si los resultados extraídos mediante dos pruebas se parecen lo suficiente, respecto al rango normal de valores esperados para la habilidad medida, como para poder usarlas indistintamente o, por el contrario, tienden a aportar resultados significativamente diferentes.

Mientras que en alguna de las habilidades binoculares estudiadas el grado de concordancia entre varias de las pruebas fue bastante bueno, por ejemplo en la medida de los rangos de vergencia fusional negativa y entre algunas técnicas de medida de desviación ocular. En otras cualidades visuales, el grado de concordancia no fue tan alto como el que cabría esperar para pruebas que supuestamente miden la misma función, por ejemplo, en la medida de la amplitud de acomodación, de los rangos de vergencia fusional positiva o de la estereopsis. Ello lleva a pensar que aunque todos los tests de cada grupo midan la misma función subyacente, que se nombra como amplitud de acomodación, desviación ocular, estereoagudeza, etc.; hay una serie de variables superimpuestas asociadas al propio diseño de la prueba que afectan al resultado de la medida de diferente forma o en distinto grado de modo que la entidad final medida sea diferente.

Tabla 4. Concordancia entre las pruebas que miden la misma habilidad visual

HABILIDAD ESTUDIADA	DM±CDC	
AMPLITUD DE ACOMODACIÓN <ul style="list-style-type: none"> • Acercamiento – Alejamiento • Lentes negativas – Alejamiento • Lentes negativas – Acercamiento 	+1,82±4,73 D -2,94±4,50 D -4,77±5,65 D	
RETRASO ACOMODATIVO <ul style="list-style-type: none"> • Nott - MEM • Nott - CCF • Nott - AutoRx • MEM - CCF • MEM - AutoRx • CCF - AutoRx 	+0,23±0,64 D +0,72±0,95 D -0,08±0,98 D +0,52±1,00 D -0,31±0,98 D -0,82±0,92 D	
DESVIACIÓN OCULAR HORIZONTAL <ul style="list-style-type: none"> • Cover (Neutral.) - Cover (Punto central) • Cover (Neutralización) - von Graefe • Cover (Neutralización) - Maddox • Cover (Neutralización) - Thorington • Cover (Punto central) - von Graefe • Cover (Punto central) - Maddox • Cover (Punto central) - Thorington • Von Graefe - Maddox • Von Graefe - Thorington • Maddox - Thorington 	LEJOS 1,86±2,49 Δ -0,59±4,35 Δ -3,22±5,81 Δ --- -2,45±3,91 Δ -5,08±5,28 Δ --- -2,60±4,66 Δ --- ---	CERCA 2,00±2,97 Δ 0,72±6,54 Δ 0,71±7,50 Δ -0,59±5,82 Δ -1,27±6,69 Δ -1,28±7,19 Δ -2,58±6,13 Δ -0,01±6,41 Δ -1,31±4,95 Δ -1,30±5,64 Δ
DESVIACIÓN OCULAR VERTICAL <ul style="list-style-type: none"> • Von Graefe - Maddox • Von Graefe – Thorington • Maddox - Thorington 	CERCA 0,09±1,00 Δ -0,05±0,50 Δ -0,04±0,75 Δ	
RESERVAS FUSIONALES <ul style="list-style-type: none"> • Diasporámetro – Barra (LEJOS) <ul style="list-style-type: none"> ○ VFN: E / Ro / Re ○ VFP: E / Ro / Re • Diasporámetro – Barra (CERCA) <ul style="list-style-type: none"> ○ VFN: E / Ro / Re ○ VFP: E / Ro / Re 	--- / -1,2±3,5 / -1,0±3,7 Δ 3,0±9,5 / 1,4±10,7 / -2,6±10,5 Δ 3,7±8,0 / 3,8±6,8 / -1,5±7,0 Δ 5,1±13,2 / 0,3±11,2 / -1,9±9,0 Δ	
ESTEREOAGUDEZA <ul style="list-style-type: none"> • Frisby – TNO • Frisby – Randot • Frisby – Titmus • TNO – Randot • TNO – Titmus • Randot - Titmus 	-42,83±100 seg arc -7,42±30 seg arc -22,34±45 seg arc 35,41±95 seg arc 20,49±80 seg arc -14,92±30 seg arc	
Clave: DM = diferencia media. CDC = coeficiente de concordancia, D = dioptría. Δ = dioptría prismática. cpm = ciclos por minuto. Nott = retinoscopia Nott, MEM = retinoscopia MEM, AutoRx = autorrefractómetro en cerca, CCF = cilindros cruzados fusionados		

Estructura de la tesis

El presente trabajo consta de dos partes: una primera parte constituida por tres capítulos introductorios y una segunda parte en la que se detallan los experimentos realizados, incluyendo la discusión de los resultados.

En el capítulo 1 se establecen los objetivos perseguidos con este estudio así como un resumen de las aportaciones novedosas del mismo desde el punto de vista clínico. En el capítulo 2, se revisa la fundamentación teórica de la investigación de forma general. El capítulo 3, tiene por objeto describir de forma global el protocolo seguido en el estudio, se tratan aspectos como: caracterización de la muestra, selección de pruebas y orden de aplicación de las mismas y el tratamiento estadístico de los datos. Aspectos como antecedentes y estado actual del tema y metodología de la investigación se completan y particularizan para cada grupo de pruebas estudiadas dentro de los capítulos 4 al 11.

Los capítulos del 4 al 11 exponen los distintos experimentos llevados a cabo. En todos estos capítulos se sigue una estructura común, equivalente a la utilizada en los artículos científicos y se compone de los siguientes apartados: *Introducción*, *Método*, *Resultados*, *Discusión* y *Conclusiones*. Esta forma de presentar la parte empírica del estudio nos permite centrarnos en cada uno de los aspectos de la visión binocular estudiados: Los capítulos 4 al 7 tratan el estudio del sistema acomodativo: el capítulo 4 se destina a la amplitud de acomodación, el capítulo 5 a la flexibilidad acomodativa, el capítulo 6 a la respuesta acomodativa y el capítulo 7 a la acomodación relativa. Los capítulos 8 al 10 estudian el sistema de vergencias: el capítulo 8 se destina a las pruebas para evaluar la desviación ocular, el capítulo 9 a la medida de las reservas fusionales y el capítulo 10 a la flexibilidad de vergencias. Para finalizar, el capítulo 11

evalúa las pruebas que permiten estudiar el nivel superior de la visión binocular, la estereoagudeza.

Finalmente, en el capítulo 12 se incluyen las conclusiones generales de la presente tesis así como diversas propuestas para futuros estudios sobre el tema.

Para terminar, se han incorporado una serie de anexos. Los tres primeros muestran las fichas de recogida de datos y otros documentos utilizados en la investigación. A continuación, se incluyen dos anexos más en los que se hace un breve repaso de las técnicas estadísticas relacionadas con los estudios de concordancia. Los dos últimos anexos recogen las contribuciones científicas y los proyectos de investigación relacionados con el desarrollo del doctorado.

Fundamentación teórica

La medición constituye una actividad omnipresente tanto en la práctica clínica como en la investigación clínica y la calidad de las mediciones condiciona no sólo la calidad de la investigación, sino también la calidad de las decisiones clínicas que se apoyan en dichas mediciones². Optometría es medición de estructuras y funciones oculares y visuales. La optometría clínica realiza de manera habitual este tipo de mediciones con el propósito de averiguar las alteraciones de esas estructuras y/o funciones que están en la base de los problemas consultados por los pacientes. El **objetivo último de las mediciones** o pruebas optométricas es el de proporcionar los datos que permitan diagnosticar correctamente las distintas anomalías oculares y visuales, de manera que se pueda orientar eficazmente su tratamiento.

El problema de la medida es común a toda la ciencia, pero presenta características diferenciales en cada una de las disciplinas. Las disciplinas clínicas presentan unas características especiales en relación con el proceso de medición, puesto que una parte de las medidas están realizadas mediante observaciones personales, mientras que otras se realizan mediante instrumentos. El problema de la calidad de la medición clínica es complejo. Por un lado, las condiciones en las que se realiza la medición no suelen ser perfectamente controlables (variabilidad fisiológica, falta de colaboración del paciente, etc.). En segundo lugar, el proceso de medición puede verse muy influenciado por la subjetividad del observador. Finalmente, con frecuencia no existe un patrón de referencia aceptable para comprobar la exactitud de la medición. Estas dificultades añadidas explican la enorme variabilidad que presentan las mediciones clínicas².

Los objetivos de las medidas clínicas varían dependiendo de que sean reunidas con fines de investigación o para intentar solucionar las quejas de un paciente. Los datos clínicos se recopilan a través de la anamnesis, a través de la observación externa y a partir de la aplicación de pruebas clínicas que precisan o no instrumentos. En las últimas décadas están siendo desarrolladas nuevas pruebas clínicas con mayor capacidad discriminativa potencial. Sin embargo, estas pruebas a menudo están poco documentadas con respecto al rango de normalidad de los resultados y a su repetibilidad lo que hace que la información que aportan no pueda ser usada eficazmente. Si un clínico no conoce la fiabilidad de un test, su aplicación será imprecisa porque no interpretará correctamente los resultados obtenidos con dicho procedimiento.

Es muy importante **optimizar el tiempo de examen** (muy reducido al estudiar niños) y aplicar en primer lugar aquellas pruebas que nos aporten la información más valiosa. Los resultados de dichas pruebas deben permitir descartar problemas visuales y/o seleccionar el tratamiento más adecuado para el paciente. Sin una buena medida no se podrán detectar los cambios clínicos ni el éxito o fallo de un tratamiento¹⁷. Agrupando e interpretando los datos clínicos se podrá colocar al paciente en una determinada categoría clínica, es decir, se podrá emitir un diagnóstico, establecer un pronóstico y plantear un tratamiento.

El estudio de las variaciones en los resultados de las pruebas puede ser más discriminativo que sus valores absolutos, ya que las variaciones intra-sujeto casi siempre son menores que las variaciones inter-sujetos⁵. A pesar de la importancia que tiene estudiar el cambio, hay poca normativa sobre la **significación clínica del cambio**, y cuando existe, es raro que se use. En la práctica, los clínicos dan un determinado nivel de importancia a los cambios en función de su propia intuición. Para

que una prueba sea adecuada para detectar pequeños cambios visuales, debe tener una alta repetibilidad; sin embargo, una alta repetibilidad no garantiza que una prueba sea capaz de discriminar entre sujetos normales y sujetos con alteración. La capacidad de discriminación también depende del nivel de correlación entre el resultado de la prueba y la alteración¹⁸.

La experiencia clínica de los componentes de nuestro equipo de investigación nos indica que en la optometría clínica existe la tendencia a utilizar determinadas pruebas de acuerdo a la formación de cada profesional, sin ningún criterio preciso. Como cada profesional mide una variable dada con una prueba preferida, resulta difícil comparar entre sí las distintas mediciones y aunque el nombre de la variable medida sea el mismo, lo que se está midiendo puede ser diferente, o hacerse incorrectamente. En general, **faltan criterios para elegir una u otra prueba** para medir cada variable visual. Se desconoce también, de manera general, la repetibilidad y la validez de las distintas pruebas utilizadas habitualmente. Por último, hay problemas de estandarización de la realización de cada prueba. Esta situación nos indujo a considerar la necesidad de emprender estudios de investigación tendentes a establecer las condiciones de las pruebas habitualmente utilizadas en optometría, en términos de repetibilidad y validez¹⁹.

La repetibilidad parece ser la característica más importante de una prueba, y además influye en las demás características. Por ejemplo, si la repetibilidad de un test es baja, es de esperar que los resultados obtenidos con dicho test correlacionen pobremente con las medidas obtenidas con la prueba de referencia o “*gold standard*” para la cualidad medida. Es decir, se espera que tenga una validez baja. Si un test correlaciona muy poco consigo mismo, ¿cómo va a poder correlacionar bien con otro test? Por otra parte, un test con baja repetibilidad es de esperar que aporte medidas

en un amplio rango de valores en sujetos normales. Esto hace que sea difícil que las medidas de un ojo anormal caigan fuera de este amplio rango de valores normales, es decir, es poco probable que sea un test con buena **capacidad discriminativa**¹¹. Sería importante reclamar a las compañías fabricantes que sus nuevos instrumentos se acompañaran de los resultados de estudios de repetibilidad. De este modo el clínico podría comparar la repetibilidad del nuevo test con la repetibilidad de los tests en ese momento disponibles.

En resumen, es necesario que los clínicos conozcan la **fiabilidad o repetibilidad** de los instrumentos y pruebas para determinar cómo de grande debe ser la diferencia entre distintas medidas para que signifique un cambio real en la medida. También es interesante conocer su **validez** para saber si aportan una medida real de la habilidad estudiada. La repetibilidad de una prueba o de un instrumento de medida es un prerrequisito de su validez, de manera que antes de plantearse si un instrumento mide lo que se quiere medir, hay que asegurarse de que el instrumento mide “algo” de forma reproducible, ya que si el instrumento no es repetible (fiable), resulta ocioso plantearse el problema de su validez². Por otra parte, una prueba ha de ser repetible, es decir, debe producir el mismo resultado antes de poderla comparar con otro método. Igualmente, hay que evaluar si un observador es consistente consigo mismo al aplicar una prueba antes de evaluar la concordancia entre observadores, es decir, se debe estudiar en primer lugar la repetibilidad intra-examinador antes que la repetibilidad inter-examinador³.

En este estudio se evaluó la repetibilidad intra-examinador y se compararon entre sí las distintas pruebas que evalúan la misma característica visual, estableciendo la concordancia entre ellas.

Estudios de repetibilidad de pruebas clínicas

Según Muñiz²⁰ una de las propiedades fundamentales de toda medición es su repetibilidad. Un instrumento de medida se considera repetible si las medidas que se hacen con él “carecen de errores”, es decir, son consistentes. Una prueba será repetible si cada vez que se aplica a los mismos sujetos da el mismo resultado, dentro de los márgenes de error. La repetibilidad, por tanto, se refiere al grado en que pueden reproducirse los resultados obtenidos por un procedimiento de medición. Un modo de descubrir si una medición es o no fiable consiste en repetirla o reproducirla²¹. Cuando se realiza un estudio de repetibilidad en mediciones clínicas el objetivo principal no suele ser la comprobación de que existe variabilidad, sino la identificación de las causas de las discrepancias, para intentar corregirlas².

En una medida clínica usualmente se consideran tres **fuentes potenciales de variabilidad**^{2, 22}: En primer lugar, la variabilidad debida a los sujetos en estudio, que es en definitiva lo que tratamos de medir. Para ello, se selecciona el instrumento o la prueba que nos ayude a distinguir entre los sujetos normales de los que presenten determinada alteración. En segundo lugar, hay que considerar la variabilidad debida al instrumento o procedimiento de medición. Por último, una parte de la variabilidad es atribuible al observador. La importancia relativa de estas tres fuentes de variabilidad depende del tipo de variable que se esté midiendo. Por ejemplo, la variabilidad debida al observador cobra mayor importancia en las pruebas de tipo objetivo frente a las subjetivas⁵. Lo comentaremos con más detalle:

- a) Variabilidad debida a los sujetos del estudio: Referida a la variabilidad biológica intrínseca del individuo que se da entre distintas medidas, debidas por ejemplo a fluctuaciones de humor, a los ritmos circadianos, al tiempo transcurrido desde la toma del tratamiento, cambios fisiológicos del sujeto que pueden ser resultado del

cansancio, estrés, postura, falta de atención, cambios en el estado acomodativo, etc. Para minimizar la variación asociada al sujeto, es importante que las condiciones bajo las cuales el sujeto es evaluado sean tan idénticas como sea posible. En los estudios de repetibilidad debe seleccionarse muy cuidadosamente el intervalo de tiempo. Ya que si es demasiado largo, la falta de coincidencia entre los resultados puede deberse a variaciones que tengan un significado propio (no debidas al azar), mientras que si es demasiado corto es posible que no haya dato tiempo suficiente para que aparezcan las fluctuaciones debidas al azar.

b) Variabilidad inherente al instrumento o al método con el que se realiza la medición.

Debida a fluctuaciones de los factores ambientales, como temperatura, ruido de fondo, grado de privacidad, falta de una buena interacción entre el clínico y el paciente, etc. Además, una mala calibración de la herramienta diagnóstica también puede llevar a errores de medida.

c) Variabilidad debida al observador: Se trata de variaciones debidas a que las medidas han sido realizadas por distintos examinadores, donde intervienen aspectos como la forma de dar las instrucciones al paciente y la interpretación de sus respuestas, el ángulo de observación, el grado de experiencia, la forma de utilizar el instrumento, la variabilidad biológica de los sentidos, etc. Además, la precisión aparente de una prueba también puede verse afectada cuando el observador hace la evaluación del valor de la medida conociendo el resultado de un comportamiento previo del test.

Las fuentes de variación no pueden eliminarse totalmente. Sin embargo, deben ser relativamente pequeñas comparadas con la variación de los resultados entre sujetos con y sin la alteración que se esté estudiando. Es decir, lo que cuenta para aceptar una prueba como repetible no es la magnitud del error, sino la relación entre éste y el

rango en que se mueven habitualmente las mediciones que se llevan a cabo con ella. Por tanto, se puede decir que la repetibilidad no es una propiedad intrínseca del instrumento o prueba, sino que depende de la variabilidad de la población a la que se aplique².

Entre las estrategias para mejorar la repetibilidad (y la validez) de una prueba se han propuesto las siguientes^{3, 22}.

- a) Estandarización de los métodos de medición: Elaborar un manual de instrucciones que contenga instrucciones específicas para efectuar las mediciones. Es necesario consignar por escrito de forma precisa las instrucciones sobre cómo preparar el entorno y el individuo, cómo llevar a cabo y anotar los resultados de una entrevista, como calibrar el instrumento, etc. Es imprescindible que las variables estén definidas perfectamente de forma operativa para que todos los observadores utilicen los mismos criterios incluso ante situaciones dudosas.
- b) Entrenamiento de los observadores: El entrenamiento mejora la coherencia de las técnicas de medición, en particular cuando están involucrados varios observadores.
- c) Refinamiento de los instrumentos: Por ejemplo, si es factible y es más preciso, conviene seleccionar instrumentos electrónicos y mecánicos para reducir su variabilidad. Otro aspecto interesante será aumentar la claridad de las preguntas de un cuestionario para evitar ambigüedades.
- d) Repetición: Sea cual sea el origen del error aleatorio, su impacto se reduce si se repite la medición. La principal limitación son los costes adicionales (tiempo, material) y las dificultades prácticas asociadas. Esta estrategia sólo mejora la repetibilidad, no la validez.

Elliot²³ asegura que los datos de repetibilidad son también valiosos cuando indican el tamaño del **cambio clínicamente significativo**. Esto permite al clínico saber si la variación encontrada en la medida entre dos visitas consecutivas es clínicamente significativa y representa una mejora o un empeoramiento de la alteración, o está dentro del rango de variación normal para el test. Elliot se lamenta de que muy pocos estudios han determinado el tamaño de un cambio clínicamente significativo para los tests optométricos. Cuando se considera la magnitud de un parámetro clínico, el conocimiento del grado de repetibilidad de las medidas es fundamental para interpretar el resultado de forma adecuada. Por ejemplo, varios estudios han examinado los cambios en la heteroforia resultantes del uso del ordenador. Se han detectado cambios de hasta 3Δ , y la pregunta que debemos hacernos es si estos cambios son clínicamente significativos o corresponden a una variación normal test-retest²⁴.

Respecto a la **relación entre la repetibilidad y la validez**, no hay que olvidar que una prueba puede tener una alta repetibilidad y sin embargo no ser en absoluto válida. Es decir, pueden estar libres de error aleatorio (ser repetibles) incluso aunque no estén libres de error sistemático o sesgo (no ser válidas). Por otra parte una prueba puede tener una validez alta como promedio de una serie de medidas y sin embargo, tener una repetibilidad baja debido a que las medidas se distribuyen en un amplio rango de valores, es decir, hay grandes diferencias entre algunas medidas. En resumen, repetibilidad no implica validez, ni validez implica repetibilidad. El análisis de la repetibilidad intra e inter-examinador es, en muchas ocasiones, un paso previo a la validación de una prueba diagnóstica, ya que, si hay mucha discrepancia entre sucesivas medidas realizada por el mismo examinador o si distintos observadores discrepan en los resultados, la prueba, aunque teóricamente pueda ser válida, tendrá poca utilidad clínica³.

En los últimos años se ha producido una cierta proliferación de artículos que versan sobre la repetibilidad de las pruebas optométricas. Entre las **motivaciones de los estudios de repetibilidad** se encontraban las siguientes⁴:

- a) Cuando no hay consenso sobre qué test es el más adecuado entre un grupo de candidatos que evalúan la misma faceta visual, un estudio de repetibilidad puede ayudar a seleccionar el mejor método para la práctica clínica o para su posterior uso en un estudio de investigación.
- b) Estudiar un nuevo test o instrumento que mide la misma característica ocular que otros anteriores.
- c) Dado que la variabilidad de las pruebas utilizadas en los estudios influyen en el tamaño de la muestra que debe ser estudiada²⁵ un estudio de repetibilidad nos orientará sobre el número de sujetos que deben incluirse en futuros estudios que incluyan los tests para los que se ha establecido el nivel de repetibilidad. Cuanto menos repetible sea la medida de un test, mayor deberá ser la muestra.
- d) Una falta de consenso en la utilidad clínica de un test también puede justificar la realización de un estudio de repetibilidad.

Estudios de concordancia de pruebas clínicas

Sería útil que se comparara toda técnica nueva de medición con una ya establecida. Es lógico que el optometrista clínico no disponga de tiempo para dedicarse a este tipo de estudios, pero en el ámbito de la investigación, sin duda, es un aspecto interesante y útil en el que trabajar. El propósito de estos estudios es comprobar si las dos técnicas tienen un grado suficiente de acuerdo como para que la técnica nueva pueda reemplazar a la ya existente o como para que ambas técnicas puedan usarse

indistintamente; es decir, sus resultados se parezcan tanto que ambas técnicas pueden intercambiarse²⁶. Para que pueda interesar reemplazar una técnica de medición tradicional por otra nueva, esta última debe aportar alguna ventaja respecto a la precedente: ser más sencilla de aplicar, más económica, más rápida, más eficiente, etc.

Es frecuente que aquellas pruebas clínicas de evaluación de la visión binocular, que se usen para medir la misma capacidad, se diferencien en varios aspectos asociados tanto a los estímulos de fijación utilizados, como a los elementos o metodología usada al hacer la medida. Sin embargo, a pesar de estas diferencias, en ocasiones las pruebas se usan de forma intercambiable sin considerar si miden lo mismo.

Mediante estudios de **concordancia** se evalúa si las diferencias técnicas producen resultados similares cuando se aplican al mismo sujeto y en forma simultánea o con mínimas diferencias de tiempo que garanticen la ausencia práctica de variabilidad y condiciones muy similares. De modo que los cambios en los resultados no se deban a cambios fisiopatológicos en la variable medida.

Existen varias técnicas estadísticas comúnmente utilizadas para estudiar la concordancia entre dos métodos o entre observadores. La elección de una técnica u otra depende del tipo de variable que se evalúa. Las variables correspondientes a las diferentes pruebas clínicas de visión binocular que se manejan en este estudio corresponden a variables continuas, donde los números tienen representación exacta de su valor. En el Anexo V se hace un repaso de varios procedimientos estadísticos que correcta o incorrectamente han sido habitualmente utilizados para estudiar el grado de concordancia entre dos pruebas.

Metodología general

Sujetos

TAMAÑO DE LA MUESTRA

Los estudios de repetibilidad intra-examinador y los de concordancia entre pruebas comparten los mismos supuestos y precisan de los mismos cálculos para determinar el tamaño necesario de la muestra de sujetos. El cálculo que se utilizó correspondió a la situación en que cada observador realizaba dos mediciones de cada sujeto y se deseaba evaluar el grado de acuerdo entre ambas mediciones. Se estableció el porcentaje de error (discordancia) entre ambas medidas que se esperaba encontrar en torno al 5% ($e=0,05$), se marcó en un 5% la precisión con que interesaba realizar la estimación del grado de acuerdo ($i=0,05$) y la confianza con que se quería expresar el resultado en el 95%. A continuación trasladamos estos valores a la Tabla 5 extraída del libro de Argimon y Jiménez³ que presenta los resultados para el cálculo en las situaciones más habituales y se determinó que era necesario utilizar una muestra de al menos 41 sujetos.

Tabla 5. Número de sujetos necesarios para la estimación de la concordancia intraobservador (Tabla H de la sección de anexos del libro de Argimon y Jiménez³).

e	Amplitud del intervalo de confianza del 95% (2i)					
	0,025	0,050	0,075	0,100	0,125	0,150
0,025	316	79	--	--	--	--
0,050	653	164	73	41	--	--
0,075	1017	255	113	64	41	29
0,100	1418	355	158	89	57	40
0,125	1868	467	28	117	75	52
0,150	2384	596	265	149	95	67

e: porcentaje de error o discordancia entre ambas mediciones que se espera encontrar.

CAPTACIÓN DE SUJETOS Y CRITERIOS DE INCLUSIÓN

Se decidió reclutar la muestra invitada a nuestro estudio entre la población de estudiantes de 1º curso de la Escuela de Óptica perteneciente a la Universidad Complutense de Madrid. Además de la obvia comodidad que representa contar con participantes tan próximos al lugar donde se realiza la investigación, otra razón para estudiar a esta población es que estos estudiantes no están adiestrados en la aplicación de las pruebas que se les van a aplicar al ser estudiantes noveles, situación que pudiera sesgar los resultados y no ser directamente extrapolables a otra población clínica aleatoria de este rango de edad y con demandas visuales similares.

Al contactar con los sujetos se les explicó brevemente el objetivo del estudio y su cometido dentro del mismo. Todos los participantes leyeron y firmaron un **consentimiento informado** (ver Anexo I). El protocolo del estudio cumplió los principios de la Declaración de Helsinki.

Los **criterios clínicos de inclusión** en el estudio fueron:

1. Ser alumno no repetidor de primer curso universitario y no tener aprobada o estar simultáneamente matriculado de ninguna asignatura de carácter clínico correspondiente a cursos superiores.
2. Sin historia de cirugía refractiva, estrabismo, nistagmo o ambliopía.
3. No estar tomando ningún medicamento ni padecer ninguna enfermedad de las que se conozca que pudieran afectar a la acomodación, las vergencias o la motilidad ocular.
4. No presentar signos de ninguna enfermedad ocular.

5. Presentar una agudeza visual (AV) compensada mayor o igual a 0,9 (20/22) en cada ojo tanto en lejos como en cerca.
6. No presentar alteraciones acomodativas o de vergencias. Sujetos asintomáticos, sin una tendencia de valores de las pruebas acomodativas y de vergencias que permitiera diagnosticar una alteración binocular¹⁴. No obstante, dadas sus altas demandas visuales en cerca, estos sujetos sí pertenecen a una población con riesgo de padecer dichas alteraciones binoculares. Se descartó a estos sujetos porque puede que respondieran peor a las pruebas debido a que su sistema visual no funciona correctamente.

La captación inicial de sujetos se llevó a cabo a través de cortas charlas introductorias impartidas en el último cuarto de hora de una clase correspondiente a cada uno de los cuatro grupos de alumnos de primer curso. En dicha presentación, tras destacar el interés que tendría para ellos como alumnos noveles este primer contacto con la investigación clínica propia de los estudios que estaban cursando, se comentó el objetivo de la investigación que pretendía llevarse a cabo y su papel dentro de la misma. A continuación se expusieron los requisitos principales que debían cumplir los participantes interesados (ver listado de criterios de inclusión) y se procedió a repartir el consentimiento informado (ver Anexo I) y el cuestionario de sintomatología e historia ocular (ver Anexo II) que deberían traer cumplimentado el día asignado para su primera revisión.

A esa primera cita acudieron 77 participantes, en ella se comenzaba revisando el cuestionario que el participante traía relleno con datos como edad, sexo e historia ocular. Este cuestionario permitió descartar a 9 sujetos con antecedentes de cirugía refractiva (1 sujetos), estrabismo, ambliopía o enfermedad ocular (7 sujetos) y por

estar tomando varios medicamentos que podrían afectar a su visión binocular (1 sujeto).

A continuación, se establecieron las características optométricas de cada persona mediante las siguientes pruebas:

- Agudeza visual (AV) monocular y binocular, con y sin compensación. Con esta prueba se comprobó que todos los sujetos presentaran una AV decimal de Snellen mayor o igual a 0,9 en cada ojo tanto en lejos como en cerca. Para tomar la AV en lejos se utilizó el proyector (Shin-Nippon CP-30) enfocado sobre una pantalla situada a 6 m y se usó como optotipo la E de Snellen. Para medir la AV en cerca se utilizó la tarjeta de letras de Sloan.
- Medida de la compensación óptica habitual con el frontofocómetro (Nidek LM-990).
- Queratometría y refracción objetiva, medidas con autorrefractómetro-queratómetro (Topcon KR 7000P).
- Refracción subjetiva que se llevó a cabo siguiendo la metodología habitual, usando optotipos de Snellen proyectados a 6 m (proyector Shin-Nippon CP-30) y foróptero manual (Shin-Nippon BR-7). Se graduó buscando el máximo positivo con la mejor agudeza visual, terminando siempre con un subjetivo binocular sin fusión¹⁵. La refracción obtenida fue la que se utilizó al aplicar las pruebas estudiadas.

A través de los resultados de estas pruebas iniciales se descartó a otros 4 sujetos por presentar una AV inferior a 0,9 en uno o en ambos ojos. Los 64 participantes restantes pasaron a realizar la primera tanda de medidas de las pruebas sometidas a estudio

(ver Tabla 7). Todas estas pruebas constituyen un minucioso examen de la visión binocular. Un optometrista diferente a los que se encargaban de aplicar las diferentes pruebas sujetas a estudio se encargó de analizar en conjunto los resultados de este examen optométrico y pudo determinar si los sujetos presentaba una visión binocular normal. Tras esta primera sesión de medidas se descartó a otros 2 sujetos que resultaron ser sintomáticos y en los que pudo diagnosticarse alguna alteración acomodativa o de vergencias¹⁴, en concreto un sujeto con insuficiencia de acomodación y uno con insuficiencia de convergencia. También se tuvo que descartar del estudio a un sujeto que entre la primera y la segunda sesión de medidas comenzó un tratamiento de ortoqueratología.

Tras excluir a aquellos sujetos que no cumplieron alguno de los criterios de inclusión, la investigación se llevó a cabo con 61 sujetos cuyo **rango de edad** fue de 18 a 32 años (media = 19,74; DE = 2,5 años). El rango de ametropías fue de -0.50 a -6.50 D de miopía y de +0.50 a +1.00 D de hipermetropía y hasta -2.75 D de astigmatismo. Las características principales de la muestra aparecen reflejadas en la Tabla 6.

Tabla 6. Características principales de la muestra. Género, medias \pm desviación estándar (DE) de la agudeza visual decimal equivalente de Snellen (AV) y de diferentes parámetros relacionados con las ametropías.

N = 61 (mujeres = 45, varones = 16)	Media \pm DE
Edad	19,74 \pm 2,50 años
AVcc en visión lejana OD	1,38 \pm 0,17
AVcc en visión lejana OI	1,37 \pm 0,17
AVcc binocular visión de lejana	1,43 \pm 0,14
Rx subjetiva OD (esfera)	-0,68 \pm 1,61 D
Rx subjetiva OD (cilindro)	-0,25 \pm 0,50 D
Rx subjetiva OI (esfera)	-0,58 \pm 1,53 D
Rx subjetiva OI (cilindro)	-0,25 \pm 0,51 D

Procedimiento

La Tabla 7 recoge el listado de las pruebas en las que se realizó el estudio de repetibilidad clasificadas según la habilidad que evalúan. Se trató de incorporar aquellas pruebas que más comúnmente se utilizan en la práctica clínica para evaluar la visión binocular. La única excepción a este criterio correspondió a la evaluación de la respuesta acomodativa mediante autorrefractómetro en visión próxima, donde tras revisar la bibliografía relacionada con el tema, pareció interesante incorporar esta prueba para confirmar la hipótesis de ser una técnica con buena repetibilidad²⁷.

De acuerdo con Bland y Altman²⁸, el mejor modo para estudiar la repetibilidad de un instrumento o prueba es tomar varias medidas en una serie de sujetos. Para ello se debe tener en cuenta que algunas de las características de los sujetos pueden variar con el tiempo, este es el caso de la medida del estado físico o de la salud de una persona. Cuanto más tiempo transcurra entre sesiones de medida mayor es la probabilidad de que hayan ocurrido cambios reales en su estado de salud o de que su actitud ante la prueba haya cambiado. En esta situación, la repetibilidad de la prueba será infravalorada. Esta dificultad se puede subsanar, al menos parcialmente, efectuando la segunda medición poco tiempo después de la primera. Sin embargo, si el intervalo de tiempo entre la aplicación de las dos pruebas es corto, es posible que el resultado obtenido en la segunda ocasión esté influido por el aprendizaje adquirido la primera vez, en cuyo caso ambas mediciones no serán independientes y el coeficiente de repetibilidad estará artificialmente elevado³. En este estudio se tomaron medidas en dos ocasiones separadas con un intervalo de tiempo entre ellas comprendido entre un mínimo de 24 horas (1 día) y un máximo de 10 días. Además de por razones prácticas, se estableció este intervalo de tiempo entre la sesión inicial y la sesión final porque se consideró que era un periodo suficientemente corto como para que la probabilidad de

que se produjeran cambios reales en el estado del sujeto fuera baja y suficientemente extenso como para que los resultados obtenidos en la segunda sesión de medidas no estuviera significativamente influenciado por el aprendizaje adquirido la primera vez. Salvo en unos pocos casos, lo habitual fue citarles con un intervalo de 1 semana entre sesiones de medida (media: 7,01; DE=1,60 días).

Tabla 7. Pruebas estudiadas que se clasifican según la calidad visual que miden

PRUEBAS ACOMODATIVAS	
AMPLITUD DE ACOMODACIÓN	RESPUESTA ACOMODATIVA
<ul style="list-style-type: none"> - Acercamiento - Alejamiento - Lentes negativas 	<ul style="list-style-type: none"> - Retinoscopia Nott - Retinoscopia MEM* - Autorrefractómetro en cerca - Cilindros cruzados fusionados
FLEXIBILIDAD ACOMODATIVA	ACOMODACIÓN RELATIVA
<ul style="list-style-type: none"> - Binocular con flipper - Monocular con flipper 	<ul style="list-style-type: none"> - Acomodación relativa negativa - Acomodación relativa positiva
PRUEBAS BINOCULARES	
DESVIACIÓN OCULAR	RESERVAS FUSIONALES HORIZONTALES
<ul style="list-style-type: none"> - Cover test (punto neutro) - Cover test (punto central)* - Técnica de von Graefe - Varilla de Maddox - Método de Thorington modificado 	<ul style="list-style-type: none"> - Diasporámetro - Barra de prismas - Punto próximo de convergencia
	FLEXIBILIDAD DE VERGENCIAS HORIZONTALES
ESTEREOAGUDEZA	
<ul style="list-style-type: none"> - Test de Frisby - Test TNO - Test de Randot - Test de Titmus 	
* MEM: método de estimación monocular. Punto central: media entre el punto de neutralización y el de inversión	

La repetibilidad puede ser fácil de estudiar cuando se interpreta una prueba que no requiere la presencia del paciente (ej. una retinografía o una topografía), sin embargo,

es más difícil cuando se trata de un hallazgo físico que requiere la presencia directa del paciente, ya que el recuerdo de la primera exploración puede condicionar el resultado de la segunda. Cuando en la investigación se pretende, además, comparar varias pruebas entre sí o medir la repetibilidad inter-examinador, el diseño del estudio será aun más complejo pues se debe tratar de asegurar la independencia de todas las mediciones, para lo que será muy útil usar técnicas como la aleatoriedad en la secuencia de aplicación de pruebas y las técnicas de enmascaramiento.

A la hora de aplicar las distintas pruebas estudiadas en el presente estudio se tuvieron en cuenta los siguientes **aspectos metodológicos**:

1. Excepto las medidas de estereoagudeza, por una parte, y las de desviación ocular subjetiva, por otra, que fueron aplicadas globalmente por el mismo examinador, dentro de un grupo de pruebas que evalúan la misma cualidad visual, cada prueba fue administrada por un examinador distinto (ver Anexo II) y además este examinador no tuvo acceso a las medidas de la misma variable obtenidas por los otros examinadores. Todos los examinadores fueron optometristas experimentados. El mismo examinador se encargó de aplicar en las dos sesiones las pruebas asignadas siguiendo exactamente el mismo protocolo en todos los sujetos.
2. Se establecieron escenarios para las pruebas comúnmente utilizados en la práctica clínica. Quizá la única excepción se presentó para las medidas de la estereopsis y de las flexibilidades acomodativa y de vergencias, donde los tests de fijación se situaron en atriles sobre una mesa, algo poco habitual en un gabinete optométrico.

3. Los resultados de la primera tanda de medidas no estuvieron presentes durante la segunda sesión, para evitar que el conocimiento de los mismos pudiera influir en el examinador²⁹.
4. Para reducir el efecto de las variaciones diurnas en los parámetros estudiados³⁰, todas las sesiones de medida se llevaron a cabo en el mismo rango horario, que por aspectos prácticos se situó entre las 12:30 y las 14:30.
5. Para cada sujeto todas las pruebas de cada sesión se aplicaron en el mismo día, cada sesión duraba aproximadamente 45 minutos.
6. El orden de aplicación de las pruebas se estableció de forma aleatoria* para tratar de evitar la influencia del efecto del aprendizaje o la fatiga del sujeto. El orden se asignó mediante extracción de bolas numeradas que correspondían a cada uno de los tests que evaluaban la misma cualidad visual (ver Figura 1). Cada sujeto extrajo una primera bola que indicó el primer test a aplicar, una segunda bola, para determinar el segundo, así sucesivamente hasta establecer el orden de todas las pruebas que conformaban un grupo. El Anexo II incluye las fichas utilizadas para anotar los datos de cada participante en el estudio.

A menos que se diga lo contrario, todas las pruebas de visión binocular se llevaron a cabo a través del resultado de la refracción subjetiva binocular, que correspondía al máximo positivo, con menor cantidad de cilindro que aportaba la mejor AV. La descripción detallada de la metodología seguida en la aplicación de cada prueba estudiada se ha incluido en el sub-apartado “Método” de cada uno de los experimentos.

* Como será comentado en el sub-apartado “Método” de los correspondientes experimentos el cover test, la retinoscopia Nott y el estereotest de Frisby se aplicaron antes de las restantes pruebas de su grupo.

Respecto al control de la variabilidad inherente a los instrumentos, con excepción del autorrefractómetro automático, que se controlaba que estaba calibrado al principio de cada sesión, el resto del equipamiento no era eléctrico por lo que el potencial de variación del mismo de una sesión de medida a otra era muy bajo, pudiéndose considerar que permaneció invariable.



Figura 1. Extracción de bolas numeradas para establecer aleatoriamente el orden de las pruebas

Tratamiento estadístico de los datos

Una vez obtenidos los datos de toda la muestra se procedió a su tratamiento estadístico. El análisis de los datos se realizó con el programa estadístico *Analyse-It for Microsoft Excel statistical program* (Leeds, UK. See <http://www.analyse-it.com>) y SPSS (versión 11 para Windows, SPSS Inc., Chicago, IL, USA). Tras preparar las tablas con los valores recopilados para cada sujeto en las sesiones de medida, se realizó una minuciosa revisión de las mismas en busca de valores no habituales o ilógicos, también se revisaron aquellos valores que en las gráficas parecían apartarse del comportamiento general. En cualquier caso únicamente se procedió a solventar los escasos errores de transcripción o codificación o los valores que eran claramente un error de medición o anotación, no eliminándose del análisis ningún valor extremo

(conocidos como *outliers*) que diera lugar a una vulneración de los principios de representatividad de la muestra de sujetos.

Para cada prueba estudiada se calculó el valor medio y la desviación estándar del conjunto de medidas (primera y segunda sesión) tomadas con dicha prueba. Se utilizó el **método estadístico de Bland y Altman** (ver Anexo IV) para establecer el grado de repetibilidad y el de concordancia de las pruebas^{28, 31}. Desde el punto de vista clínico, la ventaja de este método es que el acuerdo de los tests se expresa en las mismas unidades de la medida y permite al clínico establecer su propio criterio de si la diferencia es o no significativa. Se utilizó este método cuando la distribución de las diferencias fue aproximadamente normal, analizada mediante el **test de normalidad Anderson-Darling**³². Los factores calculados fueron la diferencia media (DM), la desviación estándar de las diferencias (DE), el coeficiente de repetibilidad (CDR = 1,96 x DE) y los límites de acuerdo al nivel de 95% (DM ± CDR). En los casos donde el test preliminar de normalidad fue significativo, en vez de medir el CDR como se acaba de describir, se calculó el percentil 95 del valor absoluto de las diferencias.

También se realizó la gráfica de las diferencias (eje-y) respecto a las medias (eje-x) para poder establecer los límites de acuerdo al 95% y tener una mejor idea de la repetibilidad de las distintas pruebas de forma gráfica y para facilitar la comparación de los intervalos de acuerdo de las distintas pruebas que evalúan la misma función visual. Los **límites del intervalo de acuerdo** constituyen un umbral para las diferencias de sucesivas medidas que tiene que ser superado para concluir que hay un 95% de probabilidades de que se haya producido un cambio real en el valor no atribuible a la variabilidad natural en las medidas. Los límites del intervalo de acuerdo dan una estimación de la repetibilidad de un determinado test pues delimitan el rango en el que

se espera que caigan el 95% de las diferencias entre medidas repetidas de la misma prueba realizadas al mismo individuo y bajo las mismas circunstancias.

De modo similar, al estudiar la concordancia entre pruebas se analizó la distribución de las diferencias entre las distintas pruebas que miden la misma función. Se estableció la diferencia media (DM) y el coeficiente de concordancia (CDC = $1,96 \times$ DE) o el percentil 95 del valor absoluto de las diferencias entre métodos si la distribución de las diferencias no era normal. Bland y Altman²⁸ también sugirieron que un buen sistema para comparar los resultados de dos métodos que miden la misma habilidad es la representación de las diferencias entre los resultados obtenidos con los dos métodos (eje-y) respecto a la media de los resultados de los dos métodos (eje-x) para cada sujeto. En estos gráficos se incluyen varias líneas horizontales indicando la diferencia media (DM) y los límites del intervalo de acuerdo entre los dos métodos, $DM \pm 1,96$ DE de las diferencias.

Tanto en el estudio de repetibilidad como en el de concordancia se compararon las medias mediante el **test-T** para muestras pareadas, estableciéndose el nivel de significación $p < 0,05$. No obstante a la hora de interpretar los resultados de esta prueba se tuvo en cuenta que pequeñas diferencias podían ser estadísticamente significativas pero no clínicamente significativas. Así mismo, como complemento del análisis estadístico según el método de Bland y Altman, en los experimentos en los que se comparaban más de dos pruebas se consideró adecuado realizar un **ANOVA** de dos factores (método y sesión) con medidas repetidas en ambos factores. En los casos en los que aparecieron diferencias significativas se realizaron comparaciones *post-hoc* utilizando el método de comparaciones múltiples de Scheffé para determinar la naturaleza de la significación (ver Anexo V).

PARTE II:

Estudio Empírico

Amplitud de acomodación

Introducción

Mediante la acomodación se aumenta la potencia dióptrica del ojo para conseguir enfocar en la retina las imágenes de objetos próximos. El punto próximo de acomodación y el punto lejano de acomodación son los puntos conjugados con la fovea cuando la acomodación está totalmente estimulada o totalmente relajada respectivamente; siendo la amplitud de acomodación (AA) la distancia dióptrica entre estos dos puntos. La capacidad de acomodación disminuye gradualmente con la edad llegando a ser clínicamente relevante alrededor de los 40 años, dando lugar a la presbicia. A partir de ese momento, la reserva acomodativa resulta insuficiente, lo que ocasiona que el paciente presente dificultades en las tareas en visión cercana y necesitará una adición positiva.

Actualmente, como consecuencia del aumento de las demandas visuales en visión próxima, los problemas acomodativos son una causa habitual de astenopía ocular especialmente en adolescentes y adultos jóvenes³³. En un estudio llevado a cabo con 117 pacientes menores de 35 años con síntomas astenópicos, Hokoda³⁴ encontró que las disfunciones acomodativas eran las alteraciones más frecuentes. Los síntomas suelen asociarse a la lectura y a otros trabajos de cerca y entre ellos se incluyen: borrosidad, dolores de cabeza, tensión ocular, visión doble, fatiga, dificultad al cambiar el enfoque de una distancia a otra y sensibilidad a la luz³⁵.

Una AA disminuida respecto a la esperada para la edad del paciente puede indicar la presencia de un error refractivo (sobre todo hipermetropía latente), también puede estar asociada a la presencia de varias alteraciones de tipo funcional o a una amplia

lista de alteraciones sistémicas. Además, un buen número de fármacos sistémicos y oculares pueden alterar de forma significativa la AA³⁶.

En un examen optométrico es posible evaluar la función acomodativa mediante la medida de la amplitud de acomodación, la respuesta acomodativa y la flexibilidad acomodativa. Si no se evalúa cada uno de estos tres aspectos podría pasarse por alto una alteración acomodativa^{37, 38}. Las medidas de AA informan sobre la máxima capacidad del sujeto para estimular su acomodación. Las dos técnicas más conocidas para evaluar la AA son la técnica de acercamiento y el método de las lentes negativas. Ambas técnicas requieren que el paciente identifique correctamente el momento en que el estímulo de fijación se vuelve borroso de forma mantenida. Un tercer método, la técnica de alejamiento, en la que se parte de un estímulo borroso, se espera que sea más repetible que la de acercamiento o la de lentes negativas en niños de entre 5 y 9 años³⁹. Como estos tres métodos de medida se diferencian en el punto final de la técnica, en el tamaño aparente del estímulo y/o en los componentes acomodativos estimulados se espera que su grado de repetibilidad sea distinto³⁵.

Ante las distintas técnicas optométricas disponibles para examinar la AA, la mayoría de los clínicos seleccionan un procedimiento para su rutina diaria según sus preferencias personales. Sin embargo, para determinar si se ha producido un cambio significativo en la AA, es esencial conocer la repetibilidad del procedimiento de medida y parece razonable que se utilicen los métodos que presenten una mayor repetibilidad y validez. Se han realizado varios estudios para comparar distintas técnicas de medida de la AA, pero pocos investigadores han analizado la repetibilidad de las mismas^{7, 40-42}. Por este motivo, el objetivo de este trabajo fue establecer la repetibilidad y la concordancia de los tres métodos comúnmente utilizados en la práctica optométrica

para evaluar la AA: (1) técnica de acercamiento, (2) técnica de alejamiento y (3) técnica de las lentes negativas.

Método

La AA se midió de forma monocular, tomándose medidas sólo en el ojo derecho de los sujetos mientras el otro ojo permaneció tapado con un oclisor opaco. Se utilizaron tres métodos de medida de la AA: acercamiento, alejamiento y lentes negativas. Durante la administración de estas técnicas el sujeto llevó puesta su compensación de lejos. Como estímulo de fijación se utilizó una tarjeta de cerca con un alto contraste negro sobre blanco ($\approx 80\%$). La tarjeta permaneció bien iluminada durante las medidas ya que, además de la luz de sala, se dirigió un foco de luz adicional sobre la tarjeta de fijación, controlando que no se generaran sombras sobre la misma. A continuación se describirá la metodología específica de cada técnica:

AA POR ACERCAMIENTO

Esta técnica requería que el paciente indicara al examinador el punto en que el test de fijación se volvía y permanecía borroso. Como estímulo de fijación se usó una letra correspondiente a AV unidad^ψ, situada inicialmente a unos 40 cm. La tarjeta se acercó lentamente hacia la cara del sujeto a unos 5 cm/s^δ (ver Figura 2). Para mantener la atención del sujeto se le pidió que nos indicara cuándo el estímulo empezaba a estar borroso, momento en el que se instó al sujeto a intentar volver a verlo nítido. El movimiento de la tarjeta cesó cuando el sujeto no podía identificar la letra tras 2 ó 3

^ψ Se utiliza un tamaño de estímulo cerca del umbral de resolución, pero se deja algo de "reserva de agudeza" porque utilizando estímulos dentro del umbral de resolución se puede hipoestimar la AA pues estos estímulos tan pequeños son malos estímulos acomodativos (Atchison DA, Capper EJ, McCabe KL. Critical subjective measurement of amplitude of accommodation. *Optom Vis Sci* 1994; 71: 699-706).

^δ Un movimiento demasiado rápido de la tarjeta tiende a hiperestimar la AA., mientras que si se mueve demasiado despacio (o comenzando desde muy lejos) el sujeto podría aburrirse y dar un resultado artificialmente reducido de la medida (Elliot DB. *Clinical procedures in primary eye care*. Edinburgh: Butterworth Heinemann; 2003).

segundos porque estaba demasiado borrosa. Entonces se midió (en centímetros) la distancia entre el punto de borrosidad mantenida y el plano de la gafa. Calculando la inversa de esta distancia (en metros) se obtuvo la AA en dioptrías.

AA POR ALEJAMIENTO

En esta técnica el sujeto dirigió la mirada a una tarjeta que contenía como estímulo de fijación una letra correspondiente a AV unidad. La tarjeta se alejó progresivamente a una velocidad de unos 5 cm/s desde el plano de la gafa hasta que el sujeto pudo identificar correctamente el estímulo fijado (ver Figura 2). En ese momento se midió la distancia (en centímetros) hasta el plano de la gafa. Calculando la inversa de esta distancia (en metros) se obtuvo la AA en dioptrías.



Figura 2. Medida de la amplitud de acomodación por acercamiento (izquierda) y por alejamiento (derecha)

AA CON LENTES NEGATIVAS

La técnica consiste en ir introduciendo lentes negativas en pasos de 0,25 D mientras el sujeto mira una tarjeta de optotipos situada a 40 cm (ver Figura 3). En concreto se pidió al sujeto que dirigiera su mirada hacia un optotipo perteneciente a AV 0,9. Se eligió un tamaño de optotipo ligeramente más grande al usado en las técnicas de acercamiento y alejamiento para intentar compensar, al menos en parte, la

disminución de tamaño inducida por las lentes negativas. El sujeto avisaba al examinador en el momento en que el estímulo de fijación se mantenía borroso, es decir, no podía volver a verlo nítido mediante un esfuerzo acomodativo consciente. Se anotó la lente anterior a la borrosidad mantenida. La AA se calculó como la suma de la potencia negativa adicionada, estando el sujeto emetropizado, más 2,50 D correspondientes al esfuerzo acomodativo necesario para ver nítido el estímulo a una distancia de trabajo de 40 cm.



Figura 3. Medida de la amplitud de acomodación con el método de las lentes negativas

1.1. Resultados

REPETIBILIDAD

La Figura 4 muestra las gráficas de Bland y Altman para las diferentes técnicas de medida de la amplitud de acomodación. El eje-y corresponde a la diferencia entre las medidas de ambas sesiones (Final – Inicial) y el eje-x corresponde a la media de las dos medidas (en D). Puesto que la escala del eje-y es la misma para todas las gráficas, es fácil apreciar que la técnica de las lentes negativas es la que mostró una mayor repetibilidad, pues presentó la menor diferencia media (DM = -0,08 D) y el

intervalo de acuerdo al 95% más estrecho. En estas gráficas también se puede ver que la repetibilidad no variaba en función del grado de AA alcanzada.

La Tabla 8 muestra los resultados correspondientes a la repetibilidad de las tres técnicas de medida de AA evaluadas.

Tabla 8. Repetibilidad en la medida de la amplitud de acomodación

	Media (D)	DE (D)	DM (F-I) (D)	p (test-t)	CDR (D)
AA acercamiento	13,08	2,79	-0,61	p = 0,002	±4,76*
AA alejamiento	11,25	1,77	-0,81	p = 0,06	±4,00
AA lentes negativas	8,31	1,83	-0,08	p = 0,6	±2,52

Clave: DE = desviación estándar. DM = diferencia media. CDR = coeficiente de repetibilidad, D = dioptría. F = final, I = inicial. AA = amplitud de acomodación

* En los casos donde CDR aparece con un asterisco, se ha sustituido éste por el valor del percentil 95 del valor absoluto de la diferencia

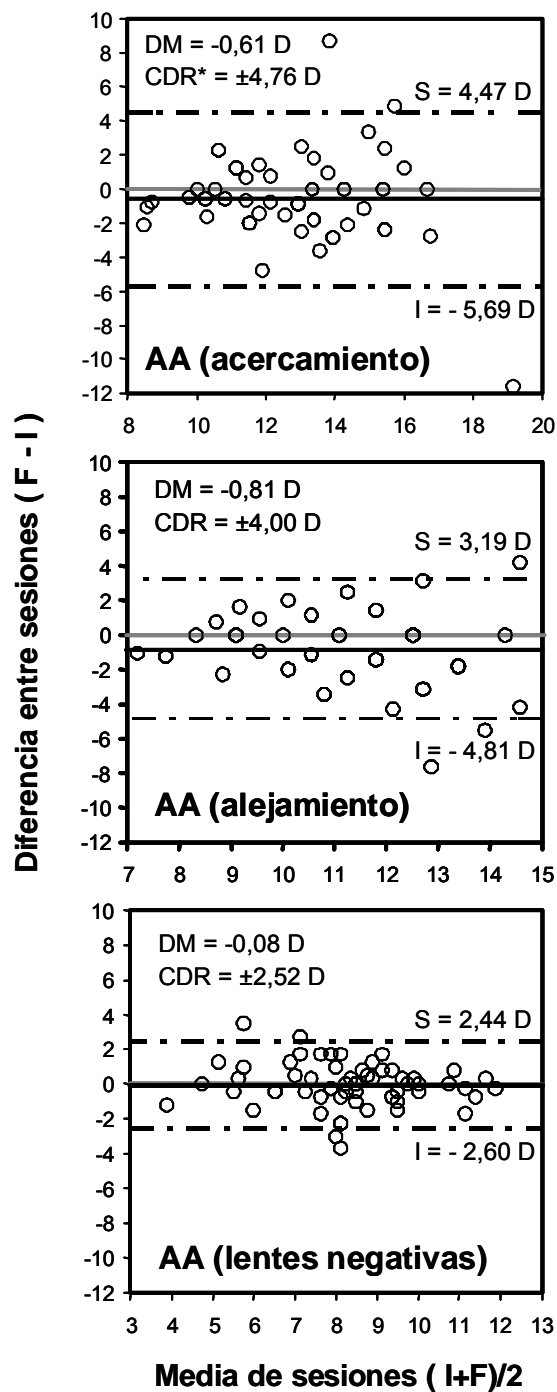


Figura 4. Diferencia media (DM), límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) y coeficiente de repetibilidad (CDR) para la medida de amplitud de acomodación (AA). En la AA por acercamiento la distribución de las diferencias no fue normal y se tomó como coeficiente de repetibilidad el percentil 95 del valor absoluto de las diferencias (CDR*).

CONCORDANCIA

La Figura 5 muestra las gráficas de Bland y Altman para el estudio de concordancia entre métodos que miden la amplitud de acomodación, el eje-x corresponde a la media entre métodos y el eje-y a su diferencia. Las líneas horizontales representadas en estas gráficas indican la diferencia media (DM) y los límites del intervalo de acuerdo al 95% entre parejas de métodos, $DM \pm 1,96 \cdot DE$ de las diferencias. En estas gráficas se aprecia fácilmente el poco acuerdo existente entre los valores de AA encontrados con el método de las lentes negativas respecto a los otros métodos. Ninguna de las gráficas ha mostrado tendencia a que la diferencia entre métodos aumente al aumentar el valor dióptrico, es decir, el grado de concordancia entre métodos es independiente del nivel de AA medida.

La Tabla 9 muestra los resultados del estudio de concordancia entre las distintas técnicas de medida de AA. Tal y como refleja la Figura 5 se puede ver que en los tres casos tanto las DMs como los CDCs fueron elevados ($\geq 4,50$ D) por lo que las tres técnicas no se pueden considerar intercambiables.

Tabla 9. Concordancia entre pruebas que miden la amplitud de acomodación (AA)

	DM (D)	p (test-t)	CDC (D)
AA acercamiento – AA alejamiento	+1,82 (acer. > alej.)	p < 0,0001	$\pm 4,73^*$
AA lentes negativas – AA alejamiento	-2,94 (alej. > lent.)	p < 0,0001	$\pm 4,50$
AA lentes negativas – AA acercamiento	-4,77 (acer. > lent.)	p < 0,0001	$\pm 5,65^*$
<p>Clave: DM = diferencia media. CDC = coeficiente de concordancia, D = dioptría. acer.= AA por acercamiento, alej. = AA por alejamiento, lent. = AA con lentes negativas</p> <p>* En los casos donde CDC aparece con un asterisco, se ha sustituido éste por el valor del percentil 95 del valor absoluto de la diferencia</p>			

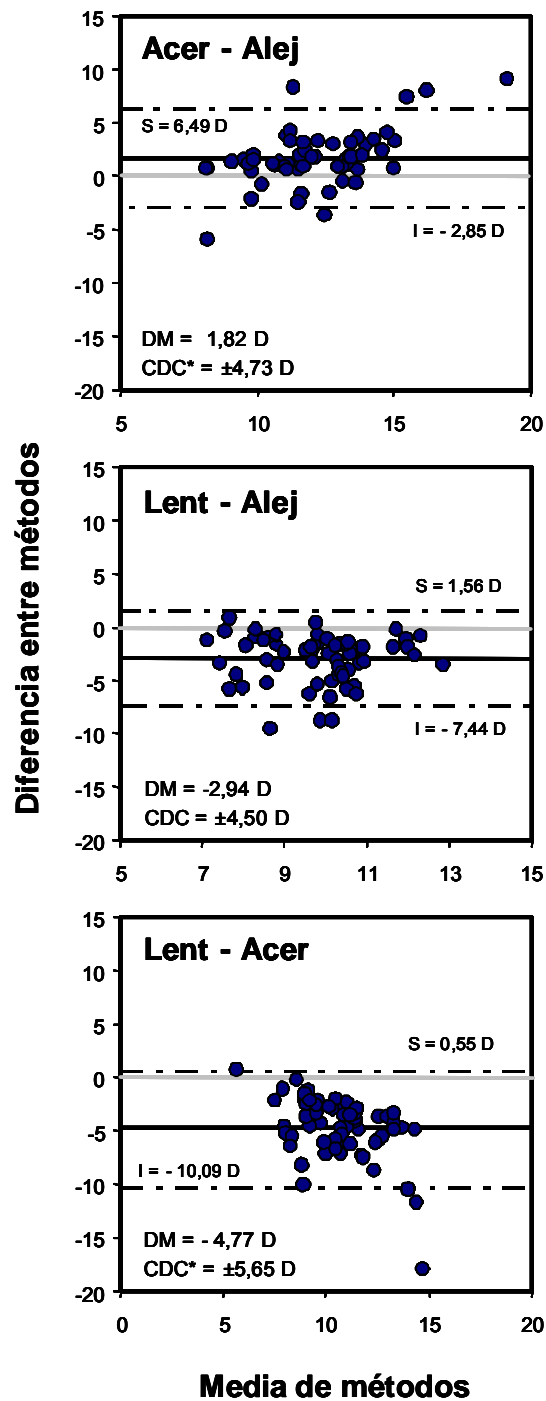


Figura 5. Gráficas de Bland y Altman para el estudio de la concordancia entre métodos que miden la amplitud de acomodación (AA). La línea continua representa la diferencia media (DM) entre las medidas obtenidas con los dos métodos. Las líneas discontinuas indican los límites superior (S) e inferior (I) del intervalo de acuerdo al 95% ($DM \pm 1,96 DE$). Cuando la distribución de las diferencias no fue normal, se tomó como coeficiente de concordancia el percentil 95 del valor absoluto de las diferencias (CDC*).

ANOVA

Los resultados de un ANOVA de medidas repetidas con dos factores (método y sesión) mostraron diferencias significativas en función del método aplicado ($F_{2,120}=110,45$; $p<0,0001$) y entre las dos sesiones de medida ($F_{1,60}=7,76$; $p=0,007$). Sin embargo, no se encontraron efectos de interacción entre ambos factores ($F_{2,120}=2,74$; $p=0,07$). El análisis *post-hoc* utilizando la técnica de Scheffé indicó la presencia de diferencias significativas entre cada uno de los valores medios obtenidos con el método de acercamiento, el de alejamiento y el de lentes negativas. La AA media obtenida con la técnica de las lentes negativas (8,31 D) fue menor que con la de alejamiento (11,25 D), que a su vez fue también menor que la obtenida con la de acercamiento (13,08 D).

1.2. Discusión

REPETIBILIDAD

Es importante conocer el CDR al medir la AA por métodos subjetivos pues establece el mínimo cambio en la AA que excede el margen de error inherente a la técnica para cada rango de edad. El menor margen de error, menor CDR, se obtuvo para la técnica de las lentes negativas donde un cambio de hasta $\pm 2,52$ D podía deberse a errores de medida y no a un cambio real en la AA.

Rosenfield y Cohen⁴¹ estudiaron la repetibilidad de las tres mismas técnicas de medida de la AA en una muestra de 13 sujetos de edad similar a nuestra población ($24,3 \pm 0,5$ años). Sus resultados mostraron que las tres técnicas presentaban un grado similar de repetibilidad con unos intervalos de acuerdo notoriamente mejores a los encontrados en nuestro caso (acercamiento: $\pm 1,44$ D, alejamiento: $\pm 1,39$ D, lentes negativas: $\pm 1,43$ D). Entre los factores que pueden contribuir a esta discrepancia está el modo en el que

han calculado la desviación estándar media de la muestra, a partir de la cual calcularon el intervalo de acuerdo al 95%. Cada sujeto fue atendido un total de 5 sesiones para cada método y en cada sesión se tomaron 3 medidas. Para cada sujeto se calculó la media y la desviación estándar de todas las medidas tomadas con el mismo método sin considerar diferentes sesiones. Después se calculó el intervalo de acuerdo para cada técnica a partir de la media de todas las desviaciones estándar individuales. Es de esperar que esta metodología permita filtrar las medidas extremas. Por otra parte, utilizaron una muestra de sujetos escasa (13 sujetos), lo que disminuye la potencia estadística de sus resultados.

Chen *et al.*⁴² aplicando el método de alejamiento en dos ocasiones separadas en una muestra de 18 adultos de entre 18 y 19 años. Utilizaron como estímulos de fijación símbolos LEA (casa, círculo, cuadrado y manzana) de tamaño equivalente a una AV decimal de aproximadamente 0,4. Encontraron una diferencia media y un intervalo de acuerdo al 95% de $0,07 \pm 1,24$ D. Las notorias diferencias entre los tests de fijación utilizados hacen que sus resultados no sean comparables a los de este estudio.

Rouse *et al.*⁷ estudiaron la repetibilidad del método de acercamiento para medir la AA en 20 niños de entre 9 y 10 años y encontraron una repetibilidad intra-examinador entre sesiones diferente para los dos examinadores ($\pm 5,32$ D y $\pm 10,48$ D). Nuestros resultados (CDR = $\pm 4,76$ D) se acercan más a los obtenidos por este autor aunque se debe tener en cuenta que en su trabajo usaron una muestra de sujetos más jóvenes en la que cabe esperar que el CDR sea mayor debido a que en niños la AA es más elevada.

CONCORDANCIA

Respecto al grado de concordancia entre métodos, nuestros resultados han ratificado los de otros trabajos previos que indicaban que la técnica de las lentes negativas suele

arrojar valores más bajos de AA que las otras dos técnicas (ver Tabla 9). Esta diferencia se ha atribuido a la disminución de tamaño del estímulo inducido por las lentes negativas⁹ que hace que aumente la demanda acomodativa⁴³. Por otra parte, la ausencia de pistas proximales en la técnica de las lentes negativas podría contribuir a la menor AA obtenida con este método. Por ejemplo, incluso aunque el ojo izquierdo del sujeto esté ocluido, los cambios en la distancia física de la tarjeta podría inducir convergencia proximal, que está ausente en la técnica de las lentes negativas. Esta vergencia proximal podría contribuir a la AA vía el cociente CA/A, además de la acomodación proximal que se haya estimulado directamente³³.

Otro aspecto a comentar es la influencia de que sea el paciente o el examinador quien se encargue de mover la tarjeta al medir la AA. En este estudio fue siempre el examinador quien desplazó la tarjeta, por lo que no se puede valorar la influencia de este factor; sin embargo, Fitch⁴⁴ demostró que en sujetos menores de 50 años se producía un incremento significativo en la AA medida cuando era el propio paciente quien se encargaba de mover la tarjeta, probablemente debido a una mayor estimulación de la acomodación proximal³³. Como ya se ha comentado, en el método de las lentes negativas la tarjeta permanecía fija a 40 cm de modo que el estímulo de la acomodación proximal permanecía relativamente constante.

Según el estudio presente, se espera que la técnica de alejamiento muestre como promedio una AA 2,94 D más alta que la técnica de las lentes negativas y la técnica de acercamiento una AA 4,77 D más alta que la de las lentes negativas. En diversos estudios se ha comparado la AA media obtenida mediante la técnica de acercamiento y la de lentes negativas, siendo esta última siempre más baja, aunque no tan baja

⁹ Según Atchison la variación del tamaño angular no es excesiva, por ejemplo, una lente de -10 D produce una disminución del tamaño aparente del estímulo menor del 20% (Atchison DA, Capper EJ, McCabe KL. Critical subjective measurement of amplitude of accommodation. *Optom Vis Sci* 1994; 71: 699-706).

como la obtenida en el estudio presente⁴⁵. Por ejemplo, Rambo y Sangal⁴⁶ obtuvieron una diferencia de 1,49 D para una población cuya edad media era 27,5 años, mientras que Kragha⁴⁷ obtuvo una diferencia de 1,23 D para una población de 23 a 27 años.

Atchison *et al.*⁴³ afirmaron que la técnica de acercamiento y la de alejamiento nos aportan una medida clínicamente más relevante que la medida de la AA con lentes negativas pues nos informan directamente del punto próximo de visión nítida. No obstante, en la técnica de acercamiento cuando se utiliza un estímulo de tamaño fijo se tiende a hiperestimar la AA debido, por una parte, al aumento del tamaño aparente del optotipo que puede retrasar la detección de borrosidad y por otra, a la mayor estimulación de la acomodación proximal³³. Esta hiperestimación de la AA es más marcada en sujetos jóvenes, con alta AA que en presbitas y pre-presbitas. Además, se debe tener en cuenta que investigaciones previas han reflejado que los métodos subjetivos de medida de la AA tienden a hiperestimar la acomodación real medida objetivamente mediante autorrefractómetros clínicos. Esta hiperestimación posiblemente se deba a que en los métodos objetivos, poco utilizados en la práctica clínica, se excluyen los efectos de factores subjetivos como la profundidad de foco^{48, 49}.

La obtención de valores más elevados de AA cuando ésta se mide con la técnica de acercamiento en comparación con la técnica de alejamiento y la de las lentes negativas ya fue indicado en investigaciones previas^{44, 50}. Tras comparar el método de acercamiento con estímulo único con otros métodos se puede considerar que la hiperestimación de la AA con este método está entre 1,50 y 2,50 D^{33, 35, 43, 51}. Este resultado no tiene implicaciones clínicas importantes para sujetos con alta AA, pues no es muy relevante que una AA de 8 ó 10 D se hiperestime en algunas dioptrías; sin embargo, en presbitas jóvenes o pre-presbitas (35 - 45 años) una hiperestimación

puede conducir a prescribir una adición de cerca menor de la adecuada o a no prescribir adición en un sujeto que sí la necesitaría.

Comparando las técnicas de acercamiento y alejamiento, cabe esperar que la técnica de alejamiento tienda a mostrar valores más bajos de AA que la de acercamiento, posiblemente debido a que en la de acercamiento el sujeto conoce la letra del test de fijación desde el principio y le será más fácil identificarla cuando se vuelva borrosa. Por el contrario, en la técnica de alejamiento el sujeto no conoce la letra y esperará hasta reconocer claramente la letra antes de responder. Por otra parte, en la técnica de alejamiento al comenzar el movimiento de la tarjeta varias dioptrías más cerca del punto próximo de acomodación, el sujeto puede relajar su respuesta acomodativa de modo que necesitará un notorio incremento de su respuesta acomodativa cuando el estímulo de acomodación decreciente caiga dentro de su rango acomodativo. Es de esperar que este cambio se produzca rápidamente en niños, pero que sea más lento en sujetos de mayor edad¹³. Estos efectos se mostraron en nuestros resultados, siendo como promedio la AA por acercamiento 1,82 D más alta que la AA por alejamiento. Chen *et al.*⁴² utilizando como estímulos de fijación símbolos LEA (casa, círculo, cuadrado y manzana) de tamaño equivalente a una AV decimal de aproximadamente 0,4 en una población de 27 sujetos de 9 a 39 años (media:19,53 años). Encontraron una diferencia media entre ambas técnicas de 1,40 D (acercamiento > alejamiento), similar a la encontrada en el presente trabajo a pesar de las notorias diferencias en el estímulo de fijación y en el rango de edad de la muestra de sujetos.

1.3. Conclusiones

La AA medida con el método de las lentes negativas presentó la mejor repetibilidad, con la menor diferencia media (-0,08 D) y el menor intervalo de acuerdo al 95% ($\pm 2,52$

D), siendo éste último notoriamente peor para las técnicas de acercamiento ($\pm 4,76$ D) y de alejamiento ($\pm 4,00$ D).

Los resultados reflejaron que la técnica de alejamiento presentaba como promedio una AA 2,94 D más alta que la técnica de las lentes negativas y que la técnica de acercamiento una AA 4,77 D más alta que la de las lentes negativas. El grado de concordancia entre las tres técnicas fue pobre lo que indica que estas técnicas no son intercambiables.

Flexibilidad acomodativa

Introducción

La flexibilidad acomodativa (FA) monocular hace referencia a la capacidad del sistema acomodativo para realizar cambios rápidos del nivel de acomodación puesto en juego. Si la prueba se realiza de forma binocular, además de evaluar la capacidad del sistema acomodativo para cambiar el enfoque rápidamente, también se estudia la interacción entre el sistema acomodativo y el de vergencias. Por ello se esperan mayores valores de flexibilidad acomodativa monocular que binocular. La medida de la FA se puede llevar a cabo, o bien alternando la potencia de las lentes interpuestas mientras el sujeto mira un estímulo acomodativo en cerca (flipper de lentes) o alternando la distancia de fijación entre dos estímulos, uno en lejos y otro en cerca. Este último método tiende a dar mayores rangos de FA debido a que en la medida con lentes suele usarse un flipper con $\pm 2,00$ D, es decir, con un salto dióptrico de 4 D; mientras que la alternancia de distancia de fijación suele hacerse desde visión lejana (6 m) a 40 cm, de modo que el salto acomodativo es sólo de $2,50$ D⁵².

La inflexibilidad acomodativa es un problema común entre los no presbíteros, que genera síntomas tales como dolor de cabeza y visión borrosa¹⁴. En la literatura no hay unanimidad en cuanto a los valores esperados de FA, no siempre están clasificados por grupos de edad y típicamente se han establecido para muestras no seleccionadas, es decir, que incluían sujetos normales y sujetos con síntomas asociados a problemas acomodativos y/o de vergencias^{53, 54}, lo que hace difícil establecer los criterios de remisión si se utiliza la prueba para cribaje visual. Levine *et al.*³⁰ estudiaron la FA en una población de 93 sujetos de edades comprendidas entre 16 y 31 años y

encontraron que la variabilidad de los resultados de la medida de FA era mayor cuanto mayores eran los síntomas de los sujetos.

En el estudio de Hennessey *et al.*⁵⁵ llevado a cabo con niños (8-14 años) se encontró que los sujetos sintomáticos realizaban significativamente peor la prueba de flexibilidad acomodativa que los sujetos asintomáticos (FAM: $p=0,01$; FAB: $p=0,002$); sin embargo, este comportamiento no se ha encontrado en adultos^{30, 56}. Esta diferente respuesta entre niños y adultos probablemente se deba al uso de la misma prueba estándar de FA con lentes de $\pm 2,00$ D a 40 cm ya que se ha visto que la medida de la FA sí puede diferenciar entre adultos sintomáticos y asintomáticos si se ajusta la demanda acomodativa a la amplitud de acomodación del sujeto⁵⁷. Además, no se recomienda medir la FA en presbíteros moderados (si la amplitud de acomodación del sujeto es menor de 4,50 D) o en presbíteros absolutos⁵⁸.

Se han identificado numerosas **variables en la evaluación de la FA** que podrían contaminar los resultados. Entre ellas están^{59, 60}:

- La profundidad de foco del sujeto
- El tamaño y la posición del texto leído
- La complejidad del estímulo
- El criterio del sujeto para juzgar si el estímulo está nítido
- El tiempo de reacción del sujeto para identificar el estímulo (automaticidad)
- El tiempo de reacción del sujeto para indicar si la tarjeta está nítida
- El tiempo necesario para que el examinador cambie las lentes
- La magnificación/minificación del estímulo inducido por las lentes
- Los factores psicológicos inherentes a la prueba.

Para poder controlar los parámetros no acomodativos como el tiempo de reacción y los aspectos psicológicos, Kedzia *et al.*⁵⁹ propusieron medir dos veces la FA, especialmente si se trata de niños, una vez con flipper neutros y otra con $\pm 2,00$ D. Siderov y Johnston⁶¹ examinaron los efectos de varios parámetros acomodativos sobre la FA medida con flipper de lentes en una población de 18 a 25 años y vieron que el número de ciclos aumentaba al aumentar el tamaño del test de una AV equivalente de Snellen de 0,8 a 0,3 ($p=0,001$), al disminuir la potencia de las lentes de $\pm 2,00$ a $\pm 1,00$ D ($p<0,001$), y al disminuir la distancia de observación de 40 a 25 cm ($p=0,05$). Este estudio enfatiza la importancia de estandarizar las condiciones de la prueba cuando se quieran comparar los resultados de distintos pacientes o hacer el seguimiento de un paciente.

Aunque hay varios estudios que han estudiado la repetibilidad de la medida de FA⁸⁻¹⁰, no se ha encontrado ninguna investigación en la que se hayan analizado los datos con los métodos estadísticos actualmente recomendados. Por ello, el objetivo de este apartado fue estudiar la repetibilidad de la medida la flexibilidad acomodativa tanto monocular como binocular siguiendo esas recomendaciones.

Método

La flexibilidad acomodativa se midió con flippers de $\pm 2,00$ D. Se contabilizaron los ciclos conseguidos en 1 minuto. En primer lugar se midió la flexibilidad acomodativa binocular (FAB) y después la monocular (FAM) del OD. Durante las medidas el sujeto llevaba colocada en gafa de pruebas su refracción subjetiva de lejos (ver Figura 6). Se mantuvo encendida la iluminación de sala y se añadió un foco de luz dirigido hacia la tarjeta de cerca que estaba situada en un atril. Como estímulo de fijación se usó una fila de letras de alto contraste negro sobre blanco ($\approx 80\%$) cuyo tamaño correspondía a

una AV equivalente de Snellen de 0,8. Se tuvo especial cuidado para no crear sombras parásitas sobre la tarjeta. Se pidió al sujeto que leyera en silencio las letras y que intentara ver nítido y simple el test lo más rápidamente posible tras cada volteo del flipper. Se le comentó que sería normal que apreciara una borrosidad momentánea del test tras cada volteo y que avisara diciendo “ya” cuando las letras estuvieran suficientemente nítidas como para poder identificarlas, momento en el cual el examinador se encargó de voltear el flipper. La medida siempre se inició interponiendo las lentes de +2,00 D. En la primera tanda de medidas se dio al sujeto un periodo de prueba de 20 segundos antes de realizar la medida.



Figura 6. Medida de la flexibilidad acomodativa con flipper de $\pm 2,00$ D.

1.4. Resultados

La Tabla 10 muestra los resultados de repetibilidad en la medida de la flexibilidad acomodativa binocular y monocular. Ambas pruebas presentaron una baja repetibilidad con DMs significativas y CDRs elevados. Los resultados del estudio se muestran gráficamente en la Figura 7. El eje-y de estas gráficas corresponde a la diferencia entre las medidas de ambas sesiones (Final – Inicial) y el eje-x corresponde

a la media de las dos medidas (en cpm). Ninguna de las dos gráficas muestra una tendencia a que la diferencia aumente con el valor de flexibilidad logrado, es decir, la repetibilidad de la prueba no cambia con el número de ciclos alcanzado.

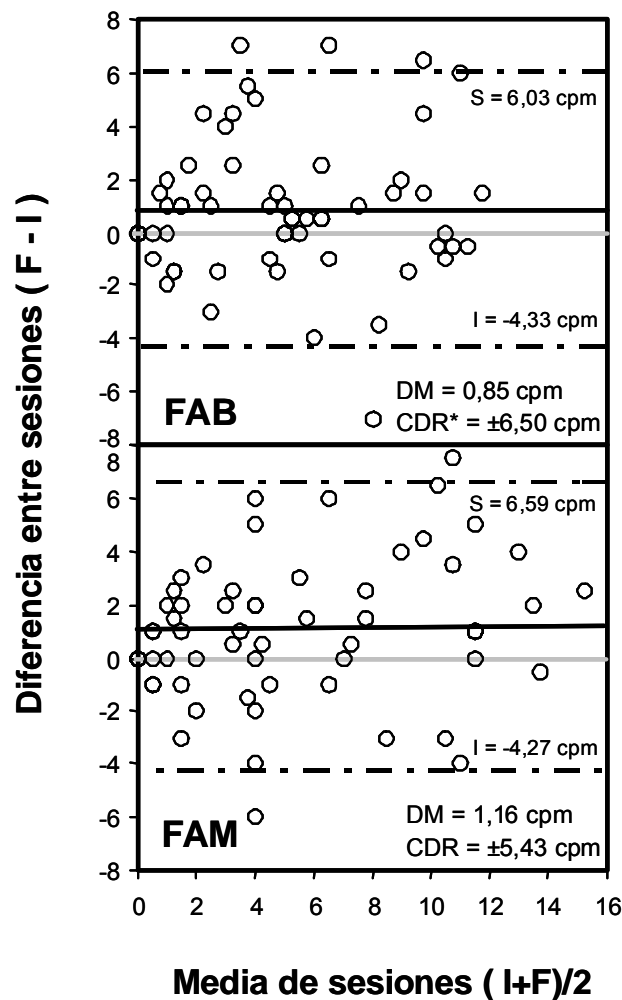


Figura 7. Diferencia media (DM), límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) y coeficiente de repetibilidad (CDR) para la medida de la flexibilidad acomodativa binocular (FAB) y monocular (FAM). En la FAB la distribución de las diferencias no fue normal y se tomó como coeficiente de repetibilidad el percentil 95 del valor absoluto de las diferencias (CDR*).

Tabla 10. Repetibilidad en la medida de flexibilidad acomodativa

	Media (cpm)	DE (cpm)	DM (F-I) (cpm)	p (test-t)	CDR (cpm)
FA binocular	4,84	3,78	0,85	0,01	$\pm 6,50^*$
FA monocular	5,71	4,89	1,16	0,002	$\pm 5,43$

Clave: FA = flexibilidad acomodativa. DE = desviación estándar. DM = diferencia media. CDR = coeficiente de repetibilidad. F = final, I = inicial. cpm = ciclos por minuto.

* En el caso donde CDR aparece con un asterisco, se ha sustituido éste por el valor del percentil 95 del valor absoluto de la diferencia

1.5. Discusión

El valor medio de FAB encontrado (4,84 cpm, DE = 3,78) se sitúa casi 3 D por debajo de los 7,72 cpm (DE = 5,15) de media del estudio de Zellers *et al.*⁶² llevado a cabo con una población de rango de edad similar. En la FAM la diferencia de nuestros resultados (5,71, DE = 4,89) con los resultados de Zellers *et al.* (11,59 cpm, DE = 5,04) aun es más marcada. Sin embargo, Scheiman *et al.*⁶³ en un estudio posterior llevado a cabo en una población más joven encontraron unos valores medios (8 a 12 años: FAB = $5,0 \pm 2,5$ cpm; FAM = $7,0 \pm 2,5$ cpm) notoriamente más bajos que Zellers *et al.*⁶² y parecidos a los de este trabajo. Es posible que en nuestro estudio las condiciones de laboratorio bajo las que se tomaron las medidas pudieran haber distraído a los sujetos contribuyendo a la baja repetibilidad de la prueba y a los bajos valores medios encontrados.

Tal y como han reflejado estudios previos^{52, 62}, la medida de FA ha sido mayor en monocular que en binocular. Este resultado es lógico porque durante la medida binocular, el cambio de vergencia acomodativa que se produce debe ser compensado por un cambio igual en la vergencia fusional opuesta. Sin embargo, en la medida

monocular sólo se cambia la demanda acomodativa, pero no la demanda de vergencia fusional.

El margen de respuesta de una población típica a la prueba de FA es muy amplio, lo que dificulta la interpretación de los resultados. Indudablemente, una de las razones para ello es que la prueba es muy subjetiva. Por otra parte, las variaciones en el diseño de las investigaciones (pruebas llevadas a cabo a distintas horas del día, toma consecutiva de tandas de medida, grupos de distinto rango de edad, etc.) hace difícil comparar los diferentes estudios. También se pueden detectar algunos fallos en el diseño del experimento como no seleccionar a una población normal (asintomática) o no enmascarar al examinador⁵³. Ello provoca que las variaciones en los resultados de distintos estudios puedan deberse a efectos de aprendizaje, variaciones diurnas, variabilidad real de la prueba y a otros factores indeterminados de modo que sea complicado extraer conclusiones generales y se desaconseja basar cualquier tratamiento únicamente en los resultados de FA⁶⁴.

Salvo un estudio no publicado realizado por Cline y Smith¹ con sólo 15 sujetos de edades comprendidas entre 8 y 30 años y con resultados poco concluyentes; los demás trabajos encontrados en los que se ha estudiado la repetibilidad de la medida de FA se llevaron a cabo en niños (de 8 a 12 años)⁸⁻¹⁰. McKenzie *et al.*⁸ midieron la FA a 66 sujetos con $\pm 2,00$ D en tres lunes consecutivos para estudiar su repetibilidad y encontraron un significativo aumento del rango de FA desde la medida inicial hasta la 2ª y 3ª sesión ($p < 0,0001$, monocular y binocular). El incremento promedio total entre la 1ª y la 2ª medidas fue 1,4 cpm (FAM) y 1,28 cpm (FAB) y entre la 1ª y la 3ª medidas fue 2,69 cpm (FAM) y 2,47 cpm (FAB). A la vista de los resultados de McKenzie *et al.*⁸

¹ Sus resultados aparecen comentados en el artículo Zellers JA, Alpert TL, Rouse MW. A review of the literature and a normative study of accommodative facility. J Am Optom Assoc 1984;55: 31-37.

y los del presente estudio con un incremento promedio de 1,16 cpm (FAM) y 0,85 cpm (FAB), se puede concluir que en la medida de la FA se manifiesta un efecto aprendizaje al repetir la prueba, comportamiento ya comentado por Levine *et al.*³⁰.

Rouse *et al.* investigaron la repetibilidad intra-sesión de la medida de FAM⁹ y FAB¹⁰ en niños con baja FA (FAM<11 cpm; FAB<8 cpm respectivamente) que fueron divididos en dos grupos: “*High fail*” (FAM: entre 6 y 11 cpm; FAB: entre 3 y 8 cpm) y “*Low fail*” (FAM<6 cpm, FAB<3 cpm). Compararon los resultados de FA en tres periodos consecutivos de 1 minuto y detectaron un leve aumento del número medio de ciclos conseguido en el 2º y en el 3º minuto respecto al número de ciclos medio conseguido en el 1º minuto. El grupo “*Low fail*” fue el que mostró menor mejoría en minutos sucesivos, así como menor variabilidad del cambio. La correlación entre la medida del 1º y 2º minuto y entre 1º y 3º minuto fue bastante buena para el grupo “*Low fail*” con coeficientes de correlación entre $r=0,67$ y $r=0,81$ y la correlación fue baja para el grupo “*High fail*” con coeficientes entre $r=0,40$ y $r=0,45$. Los autores completaron el estudio con un ANOVA de medidas repetidas con dos factores (minuto de medida y grupo de sujetos) y sólo encontraron diferencias significativas en función del minuto de medida (FAM: $p=0,0005$; FAB: $p=0,0001$) que confirma la baja repetibilidad de la prueba que también reflejó nuestro estudio.

Como en los estudios de McKenzie *et al.*⁸ y Rouse *et al.*^{9, 10}, en el estudio presente la FA mejoró significativamente en la 2ª sesión de medida respecto a los valores iniciales (FAM: DM=1,16 cpm, $p=0,002$; FAB: DM=0,85 cpm, $p=0,01$). La diferencia media encontrada junto con los amplios intervalos de acuerdo al 95% (FAM: CDR = $\pm 5,43$ cpm; FAB: CDR = $\pm 6,50$ cpm) como ya se ha comentado indican que la medida de FA no presentó una buena repetibilidad.

1.6. Conclusiones

La repetibilidad de la medida de flexibilidad acomodativa tanto monocular como binocular es pobre, tendiendo a mejorarse los valores en sucesivas tandas de medida respecto a los valores iniciales. Los clínicos deberían usar los datos de FA junto con otras pruebas diagnósticas para entender los problemas de cada paciente y plantear un tratamiento adecuado.

Respuesta acomodativa

Introducción

La medida de la respuesta acomodativa (RA) nos permite establecer el plano de enfoque del sujeto respecto al estímulo acomodativo, es decir, si se produce una hiper o una hipo-acomodación. En la práctica clínica en vez de hablar de respuesta acomodativa es frecuente trabajar con el error de acomodación, que corresponde a la diferencia entre el estímulo y la respuesta acomodativa. Se habla de retraso acomodativo si la diferencia es positiva (hipo-acomodación) y de adelanto acomodativo si la diferencia es negativa (hiper-acomodación).

La RA depende de un grupo de factores entre los que están el error refractivo, la foria y la profundidad de foco^{27, 65}. Para diagnosticar y tratar varias alteraciones acomodativas y de visión binocular es importante evaluar la RA ante un estímulo de visión próxima. Una hipo-acomodación o una hiper-acomodación respecto al plano objeto es una causa frecuente de astenopía^α. Una baja RA, es decir un retraso acomodativo mayor o igual a +1,00 D, principalmente se asocia a hipermetropía o presbicia sin compensar o hipocompensada, a insuficiencia de acomodación o a una endoforia en cerca. Sin embargo, un error acomodativo, menor o igual a 0,00 D, podría estar asociado a hipermetropía latente, a exceso de acomodación (pseudomiopía), a espasmo acomodativo o a exoforia en cerca^{11, 58}. Así mismo, una diferencia significativa entre la RA de ambos ojos puede indicar la presencia de un desequilibrio refractivo y/o acomodativo⁵⁸.

^α Cuando el error acomodativo no supere la profundidad de foco del ojo – típicamente de $\pm 0,4$ D (Campbell FW. The depth of field of the human eye. *Optica Acta* 1957;4:157-64) el estímulo de fijación puede seguir viéndose nítido, sin embargo, el paciente puede experimentar astenopía debido posiblemente al efecto de la convergencia acomodativa.

Hay varias técnicas objetivas y subjetivas para medir la RA. Entre las técnicas subjetivas una de las más utilizadas es la prueba de los cilindros cruzados fusionados (CCF), mientras que las técnicas objetivas más empleadas se basan en la retinoscopia dinámica^β, siendo las más conocidas el método de estimación monocular (retinoscopia MEM) y la retinoscopia Nott.

Se ha considerado de forma general que la retinoscopia MEM proporciona una buena medida de la RA, mostrando esta técnica una buena repetibilidad inter-examinador^{43, 66}. La hipótesis de partida es que las técnicas basadas en retinoscopia dinámica son más repetibles que las técnicas subjetivas. Además, es de esperar que sean más repetibles (especialmente en sujetos jóvenes) aquellas pruebas que no requieran la introducción de lentes adicionales para evaluar la RA, por ejemplo la retinoscopia Nott⁶⁷. Por otra parte, el autorrefractómetro objetivo de campo abierto es considerado por algunos autores²⁷ como la mejor técnica para medir la RA por lo que pareció interesante comparar las medidas de las demás pruebas estudiadas con los resultados de este instrumento.

Aunque el test de los CCF es uno de los métodos clínicos para determinar la respuesta acomodativa más comunes, parece tener una serie de deficiencias teóricas¹³. (1) El uso de estímulos rectilíneos no ayuda a situar el círculo de mínima confusión en la retina (base de la técnica) debido a su orientación horizontal y vertical. Estudios previos han demostrado que la RA ante ese tipo de estímulos varía notoriamente entre individuos⁶⁸⁻⁷⁰. (2) Además, es cuestionable que los sujetos prefieran mantener borrosos ambos grupos de líneas situando el círculo de mínima confusión sobre la retina o si preferirían alterar su RA para mejorar la nitidez de uno de

^β El principio de la retinoscopia dinámica se basa en que se observará un reflejo retinoscópico neutro cuando el punto conjugado de la retina coincida con el plano del retinoscopio.

los grupos de líneas. En el último caso el error acomodativo medido no reflejaría la respuesta habitual del paciente ante un estímulo más natural⁷⁰. (3) Por último, es de esperar que en un sujeto joven la RA cambie tras la introducción de lentes⁷¹, de modo que con la prueba de los CCF se esté evaluando también la capacidad del paciente para relajar/estimular su acomodación. Estas limitaciones hacen que en la práctica clínica, esta prueba sea más utilizada para determinar la adición tentativa para cerca que para conocer la RA¹¹.

Se han realizado bastantes estudios para comparar distintas técnicas de evaluación de la RA^{27, 66, 67, 70, 72}, pero muy pocos investigadores han analizado la repetibilidad de las mismas^{27, 73}. Por este motivo, en este estudio se quiso comparar entre sí y establecer la repetibilidad de tres técnicas objetivas de evaluación de la respuesta acomodativa (retinoscopia MEM, retinoscopia Nott y autorrefractómetro en cerca) y de una técnica subjetiva, los cilindros cruzados fusionados.

Método

Como ya se ha dicho, la RA se evaluó mediante las siguientes pruebas: (1) Retinoscopia Nott, (2) Retinoscopia MEM, (3) Cilindros cruzados fusionados y (4) autorrefractómetro en cerca. Aunque dos de estas pruebas aportan directamente la medida de la respuesta acomodativa (Nott, autorrefractómetro), en el presente estudio se decidió expresar los resultados como error acomodativo en todos los procedimientos. Como el estímulo acomodativo utilizado fue de 2,50 D en todos los casos, el error acomodativo se obtuvo como 2,50 D menos la respuesta acomodativa. Por su parte, la retinoscopia MEM y los CCF aportan directamente el resultado en forma de error acomodativo por lo que no fue necesario hacer la transformación. Un error acomodativo negativo indicaba que el sujeto hiper-acomodaba, mientras que uno positivo indicaba que estaba hipo-acomodando.

La retinoscopia Nott siempre se aplicó en primer lugar para evitar los posibles efectos de la interposición de lentes positivas o negativas sobre la RA, mientras que el orden de las restantes tres pruebas se estableció aleatoriamente mediante extracción de bolas numeradas.

Aunque sólo se anotaron las medidas para el OD, las pruebas se aplicaron en condiciones binoculares, es decir, los estímulos fueron vistos simultáneamente por ambos ojos para no eliminar el efecto de la acomodación convergente⁷⁰ aunque se ha visto que la disparidad de vergencia tiene relativamente poco efecto sobre la función acomodativa estímulo-respuesta⁵⁸, probablemente debido al bajo valor del cociente CA/A. A continuación se describe la metodología seguida en cada uno de los procedimientos.

RETINOSCOPIA NOTT

El sujeto permaneció sentado y con su compensación de lejos en el foróptero. La iluminación de sala se mantuvo atenuada y se dirigió un foco de luz hacia la tarjeta de cerca pues con bajos niveles de iluminación del test, los pacientes jóvenes pueden tender a hiper-acomodar⁷². Como estímulo de fijación se utilizó una tarjeta estrecha de visión próxima colocado a 40 cm sobre la regla graduada de VP del foróptero, situada de modo que el cero estaba alineado con el apex corneal (ver Figura 8). La tarjeta incluía varias líneas de letras de alto contraste negro sobre blanco ($\approx 80\%$). Se pidió al sujeto que mirara con ambos ojos a la línea de letras correspondiente a una AV equivalente de Snellen de 1,0. Para que la prueba informara realmente de la RA el sujeto debía esforzarse en todo momento por ver nítidas las letras⁵⁸.

Se utilizó el retinoscopio de franja Heine beta 200, que inicialmente se situó a 40 cm del paciente, justo al lado de la tarjeta de cerca[†], y se analizaron las sombras retinoscópicas del OD en el meridiano horizontal. Si en esta situación no se observaban sombras neutras, se varió la distancia de observación. Si las sombras iniciales eran inversas, el observador se acercaba hacia el sujeto hasta que observara la primera neutralización. Si las sombras iniciales eran directas, el observador debía alejarse hasta que apreciara la primera neutralización.



Figura 8. Medida de la respuesta acomodativa con retinoscopía Nott.

En el momento de obtener la neutralización, el examinador, colocando un dedo sobre la regla y anotó la distancia en centímetros desde el apex corneal hasta la posición del retinoscopio. Las distancias se anotaron con una precisión de $\pm 0,5$ cm y no se tradujeron a dioptrías hasta que se hubo evaluado a todos los sujetos. La inversa de la distancia de observación en el momento de ver sombras neutras (en metros)

[†] Al aplicar la retinoscopia Nott, aunque se intentó situar el retinoscopio lo más cerca posible de la regla, no siempre era posible estar exactamente sobre la línea de mirada del sujeto. No obstante, Brookman mostró que si el retinoscopio se encontraba a menos de 10° de la línea de mirada el error inducido era menor de 0,25 D (Brookman KE. A retinoscopic method of assessing accommodative performance of young human infants. J Am Optom Assoc 1981; 52: 865-869).

correspondía a la RA por lo que el error acomodativo se obtuvo restando al estímulo de acomodación correspondiente a 40 cm (2,50 D), la RA obtenida.

RETINOSCOPIA MEM

El sujeto permaneció sentado y portando su compensación de lejos en gafa de pruebas. La iluminación de sala se mantuvo atenuada y se dirigió un foco de luz hacia la tarjeta de cerca, igual que se hiciera con la retinoscopia Nott. Como estímulo de fijación se utilizó un test de visión próxima pegado en torno a la ventana del retinoscopio (ver Figura 9). La tarjeta de fijación incluía varias líneas de letras de alto contraste negro sobre blanco ($\approx 80\%$).



Figura 9. Medida de la respuesta acomodativa con retinoscopia MEM

Se utilizó el retinoscopio de franja Heine beta 200, que inicialmente se situó a 40 cm del plano de la gafa del sujeto controlándose la distancia mediante un cordón calibrado. Se pidió al sujeto que mirara con ambos ojos a la línea de letras correspondiente a AV unidad y que tratara de mantenerlas nítidas continuamente. En esta situación se analizaron las sombras retinoscópicas en el meridiano horizontal del

OD. Para neutralizar las sombras el examinador interponía brevemente[‡] (<0,5 s) lentes esféricas de la caja de pruebas hasta alcanzar la primera neutralización de las sombras. La lente interpuesta correspondía directamente al error acomodativo, cuando la lente era negativa significaba que el sujeto hiper-acomodaba, mientras que si era positiva estaba hipo-acomodando.

CILINDROS CRUZADOS FUSIONADOS (CCF)

Como estímulo de fijación se utilizó una tarjeta situada a 40 cm que contenía una rejilla formada por líneas horizontales y verticales de alto contraste negro sobre blanco ($\approx 80\%$). Se mantuvo encendida la luz de sala y la luz de columna. Se colocó en el foróptero la refracción subjetiva del sujeto y ante cada ojo se añadió la lente de cilindros cruzados estacionarios de la rueda de accesorios del foróptero que corresponde a unos cilindros cruzados de $\pm 0,50$ D con el eje vertical a 90° (ver Figura 10). En esta situación se preguntó al sujeto qué líneas veía más nítidas, las horizontales o las verticales. Si las veía igual de nítidas, se anotaba como resultado de la prueba (error acomodativo) 0,00 D.

Es frecuente que en la situación de partida descrita el sujeto apreciara más nítidas las líneas horizontales, es decir, que se mantuviera hipo-acomodando. En ese caso se añadieron lentes positivas en pasos de 0,25 dioptrías hasta que el sujeto apreciara igual de nítidas las líneas horizontales y las verticales, o en su defecto la primera lente con la que pasara a ver más nítidas las verticales. Por el contrario, si inicialmente veía más nítidas las líneas verticales, significaba que estaba hiper-acomodando por lo que se añadieron lentes negativas hasta obtener la igualdad. O la última lente con la que aun veía más nítidas las líneas verticales.

[‡] En esta técnica es importante interponer la lente el mínimo tiempo posible pues no se quiere que el sistema acomodativo responda al cambio sobre el estímulo acomodativo inducido por la inserción de la lente en frente del ojo.



Figura 10. Medida de la respuesta acomodativa con el método de los cilindros cruzados fusionados. (Al tomar la fotografía se dio la vuelta a la tarjeta para que mostrar el estímulo utilizado).

AUTORREFRACTÓMETRO

El autorrefractómetro Shin-Nippon SRW-5000[§] (Grand Seiko Company, Fukijama, Japón) es un autorrefractómetro binocular por infrarrojos de campo abierto que deja ver al sujeto una tarjeta real mostrada a cualquier distancia, lo que permite evaluar la RA (ver Figura 11). Según trabajos previos se ha mostrado como un instrumento válido y repetible tanto en adultos como en niños^{74, 75} y ha sido usado previamente para medir la RA con tarjetas externas⁷⁶⁻⁷⁹.

Para realizar la medida de la RA se situó la tarjeta externa en un soporte fijo a 40 cm del apex corneal del sujeto. Se utilizó una gafa de pruebas para colocar la refracción subjetiva de lejos y se pidió al participante que se esforzara en todo momento por mantener nítidas las letras correspondientes a AV unidad de la tarjeta. Se promediaron tres medidas del autorrefractómetro correspondientes al OD y se extrajo el equivalente esférico promedio. Para obtener la medida del error acomodativo, se sumó al equivalente esférico promedio la demanda acomodativa interpuesta (2,50 D) de modo

[§] También conocido como Grand Seiko WV-500 Autorrefractor

que un resultado negativo indicaba hiper-acomodación y si era positivo, hipo-acomodación.



Figura 11. Medida de la respuesta acomodativa con autorrefractómetro de campo abierto Shin-Nippon SRW-5000 con estímulo de fijación en visión próxima.

1.7. Resultados

REPETIBILIDAD

La Figura 12 muestra las gráficas de Bland y Altman para las diferentes técnicas de medida de la respuesta acomodativa, estas gráficas resultan muy útiles para comparar la amplitud del intervalo de acuerdo al 95% obtenido para los distintos métodos. El eje-y corresponde a la diferencia con signos entre las medidas de ambas sesiones (Final – Inicial) y el eje-x corresponde a la media de las dos medidas para cada sujeto (en D). Ninguna de las gráficas muestra una tendencia a que la diferencia aumente con el valor dióptrico, es decir, la repetibilidad de las pruebas no cambiaba con el valor del error acomodativo. La Tabla 11 muestra los resultados de repetibilidad de las técnicas de medida de la RA evaluadas. La mejor repetibilidad la mostró la retinoscopia Nott pues arrojó el menor intervalo de acuerdo al 95% ($CDR = \pm 0,66$ D) y una diferencia media baja ($DM = -0,10$ D).

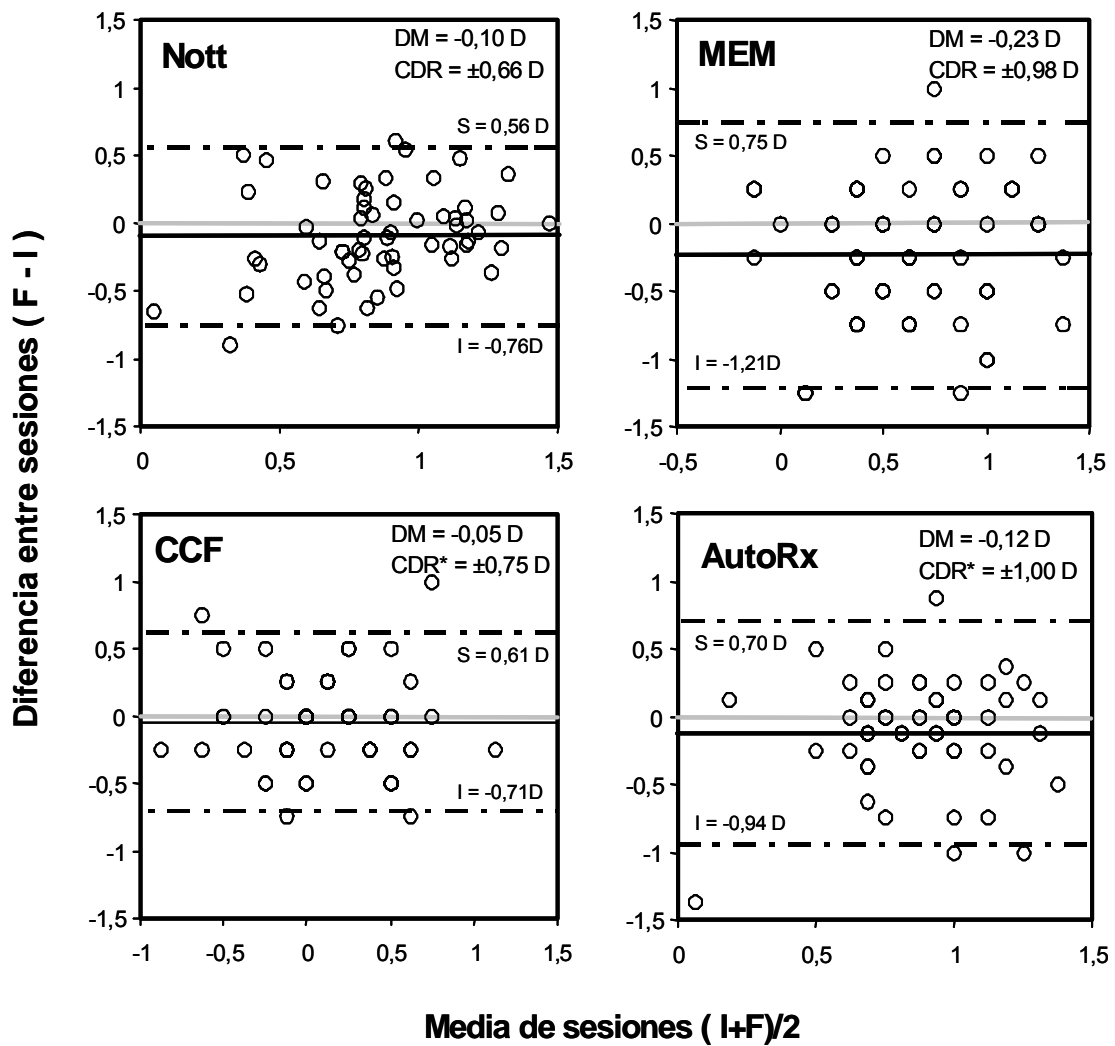


Figura 12. Diferencia media (DM), límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) y coeficiente de repetibilidad (CDR) para la medida de la respuesta acomodativa (RA). En la RA con el método de los cilindros cruzados fusionados (CCF) y con autorrefractómetro en cerca la distribución de las diferencias no fue normal y se tomó como coeficiente de repetibilidad el percentil 95 del valor absoluto de las diferencias (CDR*).

Tabla 11. Repetibilidad en la medida de la respuesta acomodativa (RA).

	Error acomodativo		DM (F-I) (D)	p(t test)	CDR (D)
	Media (D)	DE (D)			
Retinoscopía Nott	0,83	0,36	-0,10	0,02	±0,66
Retinoscopía MEM	0,63	0,37	-0,23	0,03	±0,98
CCF	0,10	0,38	-0,05	0,0005	±0,75*
Autorrefractómetro en VP	0,95	0,40	-0,12	0,3	±1,00*

Clave: DE = desviación estándar. DM = diferencia media. CDR = coeficiente de repetibilidad. F = final. I = inicial. D = dioptría. CCF = cilindros cruzados fusionados. VP = visión próxima

* En los casos donde CDR aparece con un asterisco, se ha sustituido éste por el valor del percentil 95 del valor absoluto de las diferencias

CONCORDANCIA

La Figura 13 muestra las gráficas de Bland y Altman para el estudio de concordancia entre métodos que miden la RA, el eje-x corresponde a la media entre métodos y el eje-y a su diferencia. Las líneas horizontales representadas en estas gráficas indican la diferencia media (DM) y los límites de acuerdo entre parejas de métodos, $DM \pm 1,96 \cdot DE$ de las diferencias. Dado que la escala para el eje-y es la misma para todas las gráficas, se aprecia fácilmente que no hay notorias diferencias en el grado de acuerdo existente entre los valores de error acomodativo de los distintos métodos, excepto en la gráfica que compara los dos tipos de retinoscopía dinámica, donde el intervalo de acuerdo es más estrecho, indicando que las dos pruebas son más concordantes. Por otra parte, ninguna de las gráficas ha mostrado tendencia a que la diferencia entre métodos aumente al aumentar el valor dióptrico, es decir, el grado de concordancia entre métodos es independiente del grado de respuesta acomodativa.

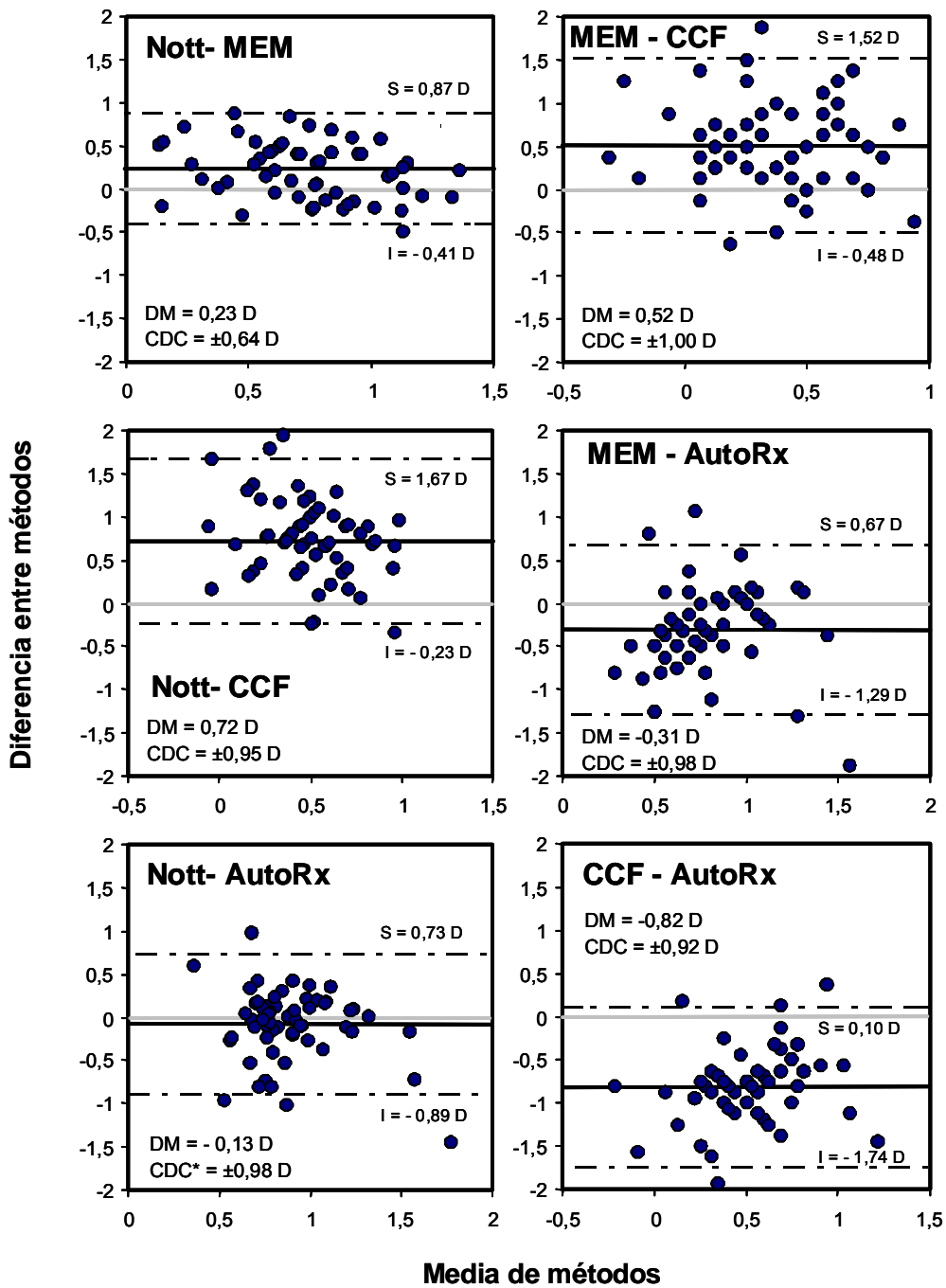


Figura 13. Gráficas de Bland y Altman para el estudio de la concordancia entre métodos que miden la respuesta acomodativa. La línea continua representa la diferencia media (DM) entre las medidas obtenidas con los distintos métodos. Las líneas discontinuas indican los límites superior (S) e inferior (I) del intervalo de acuerdo al 95% ($DM \pm 1,96 \cdot DE$). Cuando la distribución de las diferencias no fue normal, se tomó como coeficiente de concordancia el percentil 95 del valor absoluto de las diferencias (CDC*).

La Tabla 12 reúne los resultados del estudio de concordancia entre las distintas técnicas de medida de la RA. Como se puede ver en todos los casos se han obtenido intervalos de concordancia muy parecidos entre sí, pero clínicamente elevados (en torno a 1,00 D), con la única excepción del buen acuerdo que mostraron las dos técnicas de retinoscopia dinámica (CDC = $\pm 0,64$ D).

Tabla 12. Concordancia entre pruebas que miden la respuesta acomodativa (RA).

	DM (D)	p(test t)	CDC (D)
Nott - MEM	+0,23 (Nott > MEM)	0,0004	$\pm 0,64$
Nott - CCF	+0,72 (Nott > CCF)	<0,0001	$\pm 0,95$
Nott - AutoRx	-0,13 (AutoRx > Nott)	0,09	$\pm 0,98^*$
MEM - CCF	+0,52 (MEM > CCF)	<0,0001	$\pm 1,00$
MEM - AutoRx	-0,31 (AutoRx > MEM)	<0,0001	$\pm 0,98$
CCF - AutoRx	-0,82 (AutoRx > CCF)	<0,0001	$\pm 0,92$

Clave: DM = diferencia media. CDC = coeficiente de concordancia, D = dioptría. Nott = retinoscopia Nott, MEM = retinoscopia MEM, AutoRx = autorrefractómetro en cerca, CCF = cilindros cruzados fusionados

* En los casos donde CDC aparece con un asterisco, se ha sustituido éste por el valor del percentil 95 del valor absoluto de las diferencias

ANOVA

Los resultados de un ANOVA de medidas repetidas con dos factores (método y sesión) mostraron diferencias significativas en función del método aplicado ($F_{3,180}=67,31$; $p<0,0001$) y entre las dos sesiones de medida ($F_{1,60}=19,26$; $p<0,0001$). Sin embargo, no se encontraron efectos de interacción entre ambos factores ($F_{3,180}=2,53$; $p=0,06$). Un análisis *post-hoc* utilizando la técnica de Scheffé indicó que había diferencias significativas entre las medias de los distintos métodos excepto al comparar la retinoscopia Nott tanto con el autorrefractómetro como con la retinoscopia MEM.

1.8. Discusión

REPETIBILIDAD

Uno de los objetivos de este estudio fue determinar la repetibilidad de cuatro métodos de evaluación de la respuesta acomodativa. La retinoscopia Nott se ha mostrado como la técnica más repetible, donde diferencias entre medidas de hasta $\pm 0,66$ D pueden atribuirse a errores de medida. En el extremo opuesto se situaron como las técnicas menos repetibles de las estudiadas, la retinoscopia MEM y el autorrefractómetro en cerca, donde el error de medida puede llegar a $\pm 1,00$ D (ver Tabla 11). Whitefoot y Charman⁸⁰ sugirieron que a esta amplia variabilidad en los resultados de la medida de la RA podían contribuir las variaciones normales en los niveles de acomodación tónica y el efecto de la foria en cerca. Cabe esperar que el punto neutro de la retinoscopia dinámica sea más alto en endoforia que en exoforia, pues una exoforia probablemente lleve a una mayor estimulación de la acomodación convergente y consiguientemente a una mayor RA.

Se ha considerado de forma general que la retinoscopia MEM proporciona una buena medida de la RA, ya que muestra una buena repetibilidad inter-examinador⁷³. No obstante, para que la técnica sea repetible las lentes insertadas deben mantenerse colocadas sólo una fracción de segundo. Se ha visto que el tiempo de latencia de la respuesta acomodativa es de 370 mseg y la respuesta total, que incluye el periodo de latencia y los cambios de potencia del cristalino, es de aproximadamente 1 seg⁶⁶. Por ello es razonable decir que la lente de prueba no debería permanecer ante el ojo del paciente más de 1 s y deberían transcurrir un mínimo de 3 s antes de interponer una nueva lente ante el ojo del paciente⁸¹. Por otra parte, la potente luz del retinoscopio puede “cegar” momentáneamente el ojo observado, de modo que es importante que el

examinador identifique las sombras rápidamente, con uno o dos movimientos y permita de nuevo al sujeto recuperar la binocularidad⁵⁸.

Por otra parte, cuando se evalúa la RA con autorrefractómetro en sujetos amétropes es necesario medir la respuesta acomodativa a través de las lentes que compensen su ametropía. Sin embargo, ningún autorrefractómetro comercial está diseñado para este fin, ya que sería necesario tener en cuenta la potencia y la posición de las lentes compensadoras como parámetros en los cálculos del autorrefractómetro. Por ello, es posible que estos parámetros extra generen un error sistemático en las medidas de la RA con autorrefractómetro que contribuya a la baja repetibilidad obtenida en este estudio⁸².

McClelland y Saunders²⁷ estudiaron la repetibilidad de la retinoscopia Nott como método para valorar la RA en una población de entre 6 y 43 años de edad (media: 24,45, DE: 9,82 años). También estudiaron su validez, comparando los resultados con los del autorrefractómetro Shin-Nippon SRW-5000. Como en el estudio presente, la retinoscopia Nott se mostró como una técnica repetible, ya que no hubo diferencias significativas entre las dos medidas tomadas para las distintas demandas acomodativas empleadas en el estudio (10 D, 6 D y 4 D). El CDR fue $\pm 1,34$ D para una demanda de 10 D, $\pm 1,09$ D para una demanda de 6 D y $\pm 0,56$ D para una demanda de 4 D. Por otra parte, mostró ser una técnica válida, ya que los resultados presentaron un buen acuerdo con los obtenidos objetivamente mediante el autorrefractómetro, propuesto como prueba de referencia o *gold standard*. No obstante, para la demanda de 4 D, la retinoscopia Nott tendió a mostrar menor error acomodativo que el autorrefractómetro (DM=0,1 D), comportamiento que también mostraron nuestros resultados (ver Tabla 12).

CONCORDANCIA

Rosenfield *et al.*⁷⁰ evaluaron la RA para un estímulo de 2,50 D en una muestra de 24 sujetos de 25,5 a 30,1 años (media: 25,0; DE: 0,4 años). Utilizaron el autorrefractómetro Canon Autoref R-1, dos técnicas de retinoscopía dinámica (neutralización mediante lentes y por variación de la distancia de trabajo) y tres técnicas subjetivas (cilindros cruzados en cerca sin miopización y miopizando con +1,00 D y test bicromático en cerca). Al comparar los resultados del **autorrefractómetro con las otras pruebas**, la mejor concordancia lo encontraron con la retinoscopía dinámica neutralizando mediante variación de la distancia de trabajo (CDC = $\pm 0,65$ D). Para las demás técnicas la concordancia con el autorrefractómetro fue peor y similar al obtenido en el estudio presente donde los CDC estaban cercanos a $\pm 1,00$ D. En el estudio de Goss *et al.*⁶⁶ dos examinadores midieron la RA con tres técnicas (retinoscopía MEM, retinoscopía Nott y autorrefractómetro Canon Autoref R-1) a 50 sujetos (de 20 a 35 años, media 24,8). La diferencia media entre MEM y autorrefractómetro (DM=0,51 D) y entre Nott y autorrefractómetro (DM=0,51 D) fue mayor a la hallada en este estudio (Nott vs. AutoRx: DM=-0,13 D; MEM vs. AutoRx: DM=0,31 D). Sin embargo, sus coeficientes de concordancia (Nott vs. AutoRx: CDC= $\pm 0,59$ D; MEM vs. AutoRx: CDC= $\pm 0,82$ D) fueron menores que los de este trabajo que fueron de $\pm 0,98$ D en ambos casos. No obstante hay que tener en cuenta que en el estudio de Goss *et al.*⁶⁶ al promediar los valores entre sus dos examinadores pueden haberse limado los valores extremos y resultar un intervalo de acuerdo más estrecho.

Al comparar la RA de **CCF con las técnicas de retinoscopía dinámica**, Locke y Somers⁷² encontraron valores significativamente menores de error acomodativo con la prueba de los CCF que con las técnicas de retinoscopía dinámica; resultados coincidentes con los del estudio presente y los de Rosenfield *et al.*⁷⁰. Además, se vio

que no influía de manera notoria el hecho de realizar la prueba partiendo de una pequeña miopización (+1,00 D) o sin miopizar⁷⁰. En el test de los CCF la adición de lentes hace que la respuesta acomodativa no sea estática en el adulto joven, por tanto el error acomodativo no podrá ser evaluado correctamente. Por ejemplo, en la técnica de los CCF un paciente que inicialmente indica ver más nítidas las líneas horizontales puede continuar dando la misma respuesta después de la introducción de lentes positivas y la consiguiente relajación de la acomodación. En esta situación, el punto final puede estar en función no sólo de la respuesta acomodativa, sino también de la capacidad del sujeto para relajar su acomodación⁷¹.

Al comparar las dos técnicas de **retinoscopía dinámica entre sí**, Locke y Somers⁷² no encontraron diferencias significativas entre la RA obtenida a partir de la retinoscopía MEM y la Nott. Jackson y Goss⁵² también encontraron una buena concordancia entre ambos métodos en una población de 244 niños en edad escolar. Sin embargo, Cacho *et al.*⁶⁷ estudiaron la RA primero con retinoscopía Nott y después con retinoscopía MEM en 50 sujetos de 15 a 35 años (media: 23,96, DE: 3,17) y sí encontraron diferencias significativas ($p < 0,05$) (DM=0,41 D con MEM>Nott; CDC=±0,59 D). García y Cacho⁸³ realizaron otro estudio comparativo de las mismas técnicas en sujetos con anomalías binoculares y concluyeron que aunque se ha demostrado que ambas técnicas están estrechamente relacionadas, al estudiar los límites de concordancia de ambas técnicas la técnica Nott seguía siendo la más recomendada. Goss *et al.*⁶⁶ encontraron como promedio de dos examinadores una buena concordancia entre ambas técnicas con un CDC de ±0,55 D, similar al de Rosenfield *et al.*⁷⁰ de ±0,48 D y al del estudio presente (CDC=±0,64 D).

1.9. Conclusiones

Entre las diferentes pruebas estudiadas para medir la respuesta acomodativa se ha confirmado que, como en estudios previos, la mayor repetibilidad la presentó la retinoscopia Nott (prueba objetiva). Los CCF como prueba subjetiva también reflejó una buena repetibilidad, no existiendo trabajos con los que comparar este resultado.

Al estudiar la concordancia, excepto al comparar las dos técnicas de retinoscopia dinámica, en los restantes casos aparecieron diferencias notables en la RA medida con los distintos métodos con CDC próximos a $\pm 1,00$ D. MEM y Nott tendieron a dar valores de retraso acomodativo mayores (menor respuesta acomodativa) que los obtenidos con CCF. Por su parte, el autorrefractómetro fue la técnica que mostró los mayores retrasos acomodativos. Estos resultados se deben tener especialmente en cuenta a la hora de determinar la adición de cerca como tratamiento en casos de insuficiencia acomodativa o de exceso de convergencia.

Acomodación relativa

Introducción

La acomodación relativa hace referencia a la cantidad de acomodación que el sujeto es capaz de relajar (acomodación relativa negativa, ARN) o de estimular (acomodación relativa positiva, ARP) manteniendo fijo el nivel de convergencia necesario para fijar un estímulo situado a 40 cm.

Unos valores bajos de acomodación relativa pueden ser indicativo de un problema acomodativo y/o de vergencias. Una ARP baja puede representar una mala capacidad para estimular la acomodación o puede encontrarse en pacientes con endoforia o con baja vergencia fusional negativa (VFN). Una baja ARN puede estar asociada a una baja capacidad para relajar la acomodación, a exoforia o a vergencia fusional positiva (VFP) reducida⁸⁴. Una ARN mayor a +2,75 D podría estar asociada a una hipermetropía hipocompensada o a una miopía hipercompensada. Sin embargo, desde el punto de vista práctico, es raro que haya razones para medir la ARP más allá de un nivel de -2,50 D aplicando la prueba a 40 cm⁵⁸.

El objetivo de esta parte de la investigación fue estudiar la repetibilidad de la medida de la ARN y la ARP, aspecto del que no se ha encontrado resultados publicados.

Método

Como estímulo de fijación se utilizó una tarjeta que contenía una fila de letras correspondiente a una AV equivalente de Snellen de 0,8 situada a 40 cm en la barra de visión próxima del foróptero. Se colocó en el foróptero la refracción del sujeto y se ajustó la distancia interpupilar para pruebas de cerca. Se mantuvo encendida la luz de

sala y la luz de columna de modo que la tarjeta estuviera bien iluminada y con un buen contraste negro sobre blanco ($\approx 80\%$). Para medir la ARN se adicionaron binocularmente lentes positivas sobre el valor de emetropización del sujeto y para medir la ARP se adicionaron lentes negativas. El aumento de potencia siempre se realizó en pasos de 0,25 D. Se pidió al sujeto que mirara con ambos ojos el test y que se esforzara por mantener nítidas las letras en todo momento. El sujeto debía indicar el momento en el que se produjera la primera borrosidad mantenida, instante en el que notara que las letras no estaban tan nítidas y definidas como lo estaban inicialmente, incluso aunque pudiera seguir leyéndolas. La cantidad de adición positiva o negativa añadida sobre la refracción del sujeto anterior al cambio, correspondió respectivamente a la ARN o la ARP. Se midió en primer lugar la ARN y después la ARP.

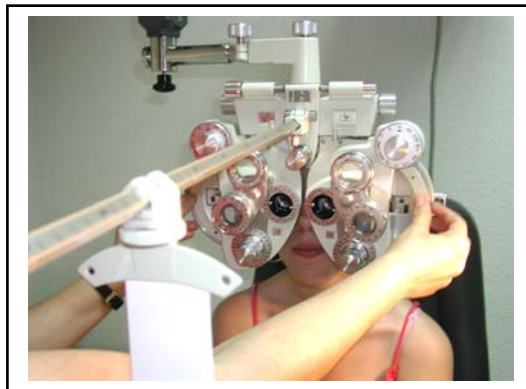


Figura 14. Medida de la acomodación relativa negativa y positiva (ARN y ARP)

1.10. Resultados

La Tabla 13 muestra los resultados de repetibilidad en la medida de la acomodación relativa negativa (ARN) y positiva (ARP). La prueba de la ARN es claramente más repetible que la de la ARP, pues tanto la DM como el CDR es notoriamente menor en la ARN. Los resultados del estudio se muestran gráficamente en la Figura 15. El eje-y

de estas gráficas corresponde a la diferencia con signos entre las medidas de ambas sesiones (Final – Inicial) y el eje-x corresponde a la media de las dos medidas para cada sujeto (en D). Ninguna de las dos gráficas muestra una tendencia a que la diferencia aumente con el valor de acomodación relativa, es decir, la repetibilidad de la prueba no cambia con el rango logrado.

Tabla 13. Repetibilidad en la medida de acomodación relativa.

	Media (D)	DE (D)	DM (F-I) (D)	p (test-t)	CDR (D)
ARN	1,88	0,53	+0,03	0,5	±0,75*
ARP	- 2,35	1,23	-0,38	0,007	±2,25*

Clave: DE = desviación estándar. DM = diferencia media. CDR = coeficiente de repetibilidad. F = final, I = inicial. D = dioptría

* En los casos donde CDR aparece con un asterisco, se ha sustituido éste por el valor del percentil 95 del valor absoluto de la diferencia

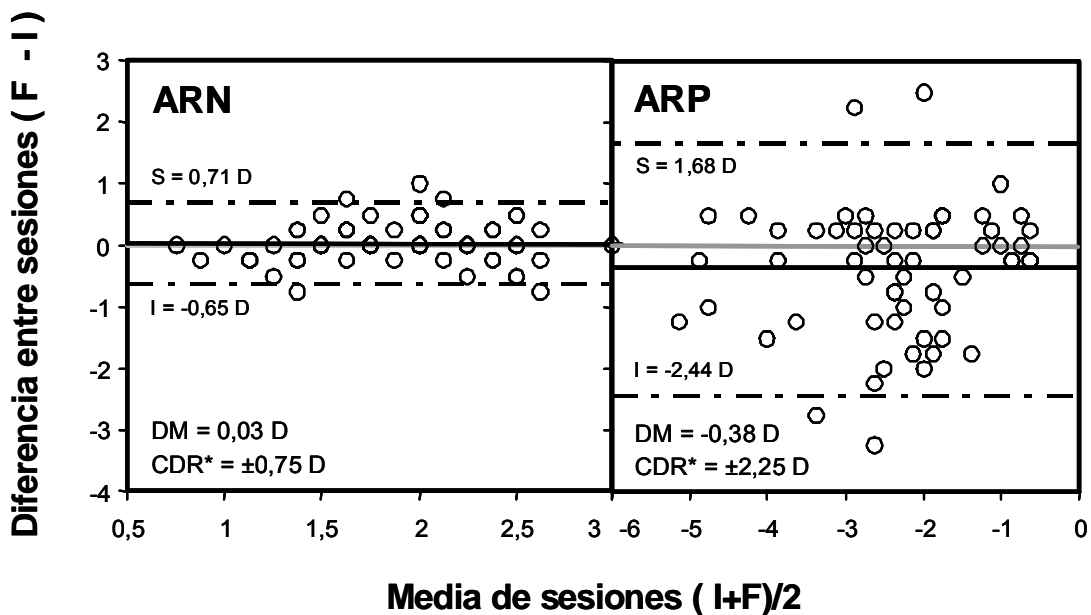


Figura 15. Diferencia media (DM) y coeficiente de repetibilidad (CDR) para la medida de la acomodación relativa negativa (ARN) y positiva (ARP). En ambos casos la distribución de las diferencias no fue normal y se tomó como coeficiente de repetibilidad el percentil 95 del valor absoluto de las diferencias (CDR*) y se anotaron los límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I).

1.11. Discusión

Goss y Zhai⁸⁵ estudiaron una población no seleccionada de 800 pre-présbitas con una amplitud de acomodación mínima de 5,00 D y obtuvieron que el valor medio de la ARN (\pm DE) para una distancia de 40 cm era $+2,00\pm 0,50$ D y que para la ARP de $-2,37\pm 1,12$ D. Estos valores son muy similares a los encontrados en el presente estudio (ARN = $+1,88\pm 0,53$ D y ARP = $-2,35\pm 1,23$ D). Es notoria la mayor variabilidad de los resultados de la medida de la ARP frente a la de la ARN.

Los resultados han mostrado una amplia diferencia entre los intervalos de acuerdo al 95% del ARN (CDR= $\pm 0,75$ D) y ARP (CDR= $\pm 2,25$ D). Mientras que la medida de la ARN es bastante repetible, la de la ARP presenta una alta variabilidad achacable a errores de medida. Entre los factores que justifiquen estos resultados posiblemente el que más influya sea la alta subjetividad de la prueba que requiere que el sujeto identifique un cambio generalmente muy sutil, especialmente en el ARP. Además, el valor de la ARP puede variar por diversos factores tales como la amplitud de acomodación del sujeto, la relación AC/A (convergencia acomodativa/acomodación) y las vergencias fusionales negativas¹⁴. Es de esperar que la variabilidad sea mayor en sujetos jóvenes con buena amplitud de acomodación que en présbitas. No se han encontrado estudios previos en los que se investigue la repetibilidad en la medida de la acomodación relativa con los que poder comparar nuestros resultados.

1.12. Conclusiones

La medida de la ARN presenta una repetibilidad bastante buena, pero la de la ARP es muy baja por lo que no se recomienda su uso aislado para la toma de decisiones clínicas.

Desviación ocular

Introducción

Para poder tener visión binocular es necesario que los ojos estén correctamente alineados y los encargados de ello son los mecanismos de fusión motora y sensorial. El diccionario de ciencias visuales define **foria o heteroforia** como la tendencia de las líneas de mirada a desviarse respecto a la posición necesaria para mantener la visión binocular única para una determinada distancia de fijación⁸⁶. Una foria es una desviación latente de los ejes visuales que se manifiesta en ausencia de estímulo de fusión. En una exoforia los ejes visuales se cruzan por detrás del estímulo de fijación, en una endoforia se cruzan por delante y en una ortoforia se cruzan sobre el estímulo de fijación. Para medir la desviación se pueden usar prismas (diasporámetro, barra de prisma o prismas sueltos) o escalas graduadas. La magnitud de la desviación se expresa en dioptrías prismáticas (Δ). La posición de foria de los ojos depende del nivel de convergencia tónica, de la respuesta acomodativa y por tanto de la convergencia acomodativa, de la cantidad de convergencia proximal, de la vergencia voluntaria y de la cantidad de adaptación de vergencias⁸⁴.

Para realizar un buen diagnóstico y tratamiento de las alteraciones de vergencias y acomodativas es importante evaluar de forma precisa la desviación que presenta el paciente tanto en visión lejana como en visión próxima. Los métodos habituales para medir la desviación ocular precisan que los ojos estén disociados. Esta situación ya está presente en el caso de tropías, pero en el caso de forias se debe eliminar la fusión.

La **disociación** se puede conseguir mediante distintas maniobras:

- Ocluyendo un ojo, como en el cover test
- Desplazando una imagen mediante un prisma, como en la técnica de von Graefe
- Distorsionando una imagen usando una varilla de Maddox, como en la técnica de Thorington
- Presentando imágenes distintas a cada ojo mediante estereoscopios.

El último sistema de disociación es más frecuente en la realización de *screening* visual, pero es menos común en la práctica clínica habitual por lo que no fue estudiado en este trabajo. Es de esperar que cuanto más tiempo permanezca disociado el paciente menor sea la adaptación de vergencias⁸⁷. Además del sistema de disociación empleado, pueden darse otras **diferencias entre los métodos de medida** de las forias como:

- El nivel de instrumentación utilizado en la medida, el nivel de iluminación, el mayor o menor control de la acomodación
- El grado de convergencia tónica puesta en juego
- El nivel de convergencia proximal inducida
- La forma de cuantificar la foria (estimación, prismas o escala graduada)
- El nivel de objetividad implicado.

Mientras que hay evidencias para poder decir que la vergencia tónica permanece relativamente estable durante las medidas⁸⁸, sin embargo, no está claro si cada prueba induce el mismo grado de convergencia proximal. Por ejemplo, se ha especulado que la linterna utilizada en la medida de la desviación con varilla de Maddox no informa correctamente de la distancia aparente del estímulo si se compara con la técnica de von Graefe o la de Thorington⁸⁹. También es importante el tipo de

tarjeta, que debería tener detalles acomodativos para generar un nivel de acomodación adecuado⁹⁰.

El cover test es una prueba objetiva^{**}, es decir, no requiere ninguna respuesta por parte del paciente; mientras que las demás pruebas que se estudiaron son de tipo subjetivo. El cover test es una prueba que se suele aplicar a todos los pacientes ya que aporta información cualitativa y cuantitativa de la desviación. Sin embargo, las pruebas subjetivas no permiten distinguir entre desviaciones latentes y manifiestas y tampoco ofrecen información de la velocidad del movimiento de refijación. Además, las pruebas subjetivas miden el ángulo subjetivo de desviación en un estrabismo y este valor puede o no corresponder con la desviación real del paciente estrábico (medida con el cover test) en función del tipo de correspondencia sensorial que presente. Incluso puede darse el caso de que la desviación subjetiva sea cero y, sin embargo, haya estrabismo, por ejemplo en una microtropía. En la práctica clínica habitual la evaluación subjetiva de la desviación ocular suele usarse más frecuentemente en sujetos no estrábicos en los que el cover test ha indicado la presencia de una foria elevada o si se ha producido un cambio significativo en la graduación del sujeto. No obstante, hay casos en los que la realización del cover test se hace difícil o imposible, por ejemplo es difícil valorar si existe o no movimiento a través de lentes de prueba de pequeño diámetro. En estos casos las pruebas subjetivas pueden hacerse especialmente útiles.

Respecto al **cover test** alternante, aunque hay poco debate sobre como llevar a cabo esta prueba, aun está sin aclarar qué valor de prisma debe ser anotado como resultado de la misma. Una posibilidad es añadir prisma hasta que no se vea movimiento (primera neutralización). Sin embargo, si se añade potencia prismática

^{**} Existe también una modalidad de cover test subjetivo basado en la apreciación por parte del paciente del movimiento *phi*.

adicional una vez que se ha llegado al valor de neutralización, el examinador puede observar que hay un rango de prismas interpuestos en los que se sigue sin observar movimiento de los ojos (rango de neutralidad). Si se siguen añadiendo prismas sobre el límite superior del rango de neutralidad se podrá observar un movimiento de los ojos contrario al inicial (punto de inversión). Por tanto, hay numerosos posibles resultados del cover test: primera neutralización, límite superior, punto de inversión y cualquier valor intermedio. La mayoría de la literatura sugiere tomar el valor de la primera neutralización, pero no hay consenso sobre qué valor es el estándar⁹¹.

La medida de la desviación ocular tanto en lejos como en cerca se encuentra dentro de las pruebas estándares dentro de la práctica optométrica. El cover test permite la medida objetiva de la desviación y ha mostrado tener una estupenda repetibilidad tanto intra-examinador como inter-examinador⁹¹. También hay bastantes métodos disponibles para medir subjetivamente la desviación ocular mediante la presentación de estímulos no fusionables a los dos ojos. Los principales estudios previos sobre medidas de forias han mostrado diferencias en la repetibilidad de estas técnicas. Por ejemplo, se ha observado que la técnica de Thorington modificada presenta un mayor nivel de repetibilidad que la medida con varilla de Maddox o la técnica de von Graefe⁹⁰.

^{92, 93}.

El **objetivo** de este estudio fue establecer la repetibilidad de los métodos de medida de la desviación ocular más utilizados en una consulta optométrica general. También se compararon los resultados de cada una de las pruebas entre sí para establecer el grado de concordancia entre los distintos métodos.

Método

Se usaron diversas técnicas de medida de la desviación ocular (ver Tabla 14): 5 técnicas de medida de la desviación horizontal en cerca (40 cm) y 4 técnicas en lejos (6 m). Dado que los niveles de acomodación o de vergencia horizontal no suelen afectar significativamente a las desviaciones verticales encontradas habitualmente en la medida de heteroforia, es suficiente realizar la medida en lejos (VL) o en cerca (VP)⁵⁸. En el estudio presente se optó por medir la desviación vertical con 3 técnicas diferentes aplicadas a 40 cm pues resulta más cómodo. Todas las medidas se realizaron en posición primaria de mirada.

Tabla 14. Pruebas de medida de la desviación ocular estudiadas.

	LEJOS	CERCA
DESVIACIÓN HORIZONTAL	<ul style="list-style-type: none"> - Prisma cover test <ul style="list-style-type: none"> - Punto de neutralización - Punto central entre neutralización e inversión - Von Graefe - Varilla de Maddox 	<ul style="list-style-type: none"> - Prisma cover test <ul style="list-style-type: none"> - Punto de neutralización - Punto central entre neutralización e inversión - Von Graefe - Varilla de Maddox - Thorington modificado
DESVIACIÓN VERTICAL		<ul style="list-style-type: none"> - Von Graefe - Varilla de Maddox - Thorington modificado

La técnica del prisma cover test fue aplicada en primer lugar por un examinador, mientras que las medidas subjetivas fueron realizadas a continuación por otro examinador que no tuvo conocimiento de los resultados del cover test. De este modo el examinador no veía influenciada su medida objetiva de la desviación por el conocimiento de los resultados de las medidas subjetivas.

El orden de aplicación de los grupos de pruebas subjetivas para diferentes distancias y componentes de la desviación fue: (1) desviación horizontal en lejos, (2) desviación

vertical en cerca y (3) desviación horizontal en cerca. Dentro de cada grupo, el orden de aplicación de las pruebas subjetivas para cada sujeto se determinó de forma aleatoria mediante extracción de bolas numeradas para asegurar que no influyera de forma importante sobre los resultados el efecto del aprendizaje o la fatiga del sujeto. No se dio ninguna indicación a los sujetos sobre si sus respuestas eran correctas o incorrectas. Se dio al menos 15 segundos de descanso entre cada método de medida durante los cuales los sujetos dirigían su mirada a lo lejos (6 m) para minimizar los efectos de la adaptación prismática⁹⁴.

Se estableció que una desviación horizontal negativa indicaba exodesviación y una positiva endodesviación. Una hiperdesviación de OD se registró como positiva y una hipodesviación de OD como negativa. Dada la naturaleza de los datos (endo era positivo y exo era negativo) fue conveniente trabajar también con la media del valor absoluto de las diferencias para que los signos negativos y positivos no se cancelaran entre sí, minimizando cualquier diferencia real que pudiera estar presente. Por ello se calcularon las diferencias y el valor absoluto de la diferencia de magnitud entre forias correspondientes a las dos sesiones de medida tal y como propusieron Johns *et al.*⁹¹.

A continuación se detalla la metodología seguida en las distintas pruebas.

MEDIDA DE LA DESVIACIÓN HORIZONTAL

El objetivo fue medir la desviación horizontal de los ojos en condiciones de disociación.

PRISMA COVER TEST OBJETIVO EN VL Y VP

Como estímulo de fijación se utilizó una única fila de letras correspondiente a un nivel de AV de 0,8. Para las medidas de lejos se proyectó a 6 m y en cerca se usó una tarjeta que fue sujeta por el paciente a 40 cm (ver Figura 16). Durante las medidas se mantuvo encendida la luz de sala y la de columna. El sujeto llevaba colocada su

refracción de lejos en gafa de pruebas con lentes de prueba de gran diámetro o utilizaba su propia gafa si la graduación era correcta. Se pidió al sujeto que se fijara en una única letra de la tarjeta y que se esforzará en mantenerla nítida durante el tiempo que durara la medición.



Figura 16. Medida de la desviación ocular con la técnica del prisma cover test en visión próxima (izquierda) y en visión lejana (derecha).

Se aplicó la prueba del cover test según la metodología habitual para determinar la dirección del movimiento¹¹. A continuación se tomó una barra de prismas con pasos de potencia prismática de 1, 2, 4-20 en pasos de 2Δ y 25-40 en pasos de 5Δ y se procedió a compensar la desviación horizontal de forma objetiva situando la barra de prismas ante el ojo derecho (OD). El primer dato que se anotó correspondió a la potencia y base del prisma que lograba eliminar el movimiento de ambos ojos al realizar un cover test alternante (primera neutralización). Posteriormente se procedió a aumentar la potencia prismática hasta que se observara por primera vez un movimiento ocular en dirección contraria al inicial (punto de inversión). En este estudio se compararon los resultados de la primera neutralización y del punto medio entre la primera neutralización y el punto de inversión. Las medidas fueron llevadas a cabo por un experto optometrista pediátrico con más de 15 años de experiencia clínica.

En aquellos sujetos en los que no se observó movimiento inicial (ortofóricos) se anotó “cero” como primera neutralización y se procedió a aumentar la potencia del prisma con base interna hasta que se observó movimiento (punto de inversión).

VON GRAEFE EN VL Y VP EN DIRECCIÓN HORIZONTAL

Como estímulo de fijación se utilizó una única columna de letras correspondiente a un nivel de AV de 0,8; para las medidas de lejos se proyectó a 6 m y en cerca se colocó la tarjeta en la barra del foróptero a 40 cm. Durante las medidas se mantuvo encendida la luz de sala y se añadió la de columna para las medidas de cerca. Se situó en el foróptero la graduación del sujeto y se ajustó adecuadamente la distancia interpupilar.



Figura 17. Punto de para la medida de forias horizontales con la técnica de von Graefe utilizando el foróptero.

Como prisma disociador se colocaron en el diasporámetro del OD 6Δ base superior, que desdoblaba el test en dos columnas de letras: la de abajo, vista por el OD y la de arriba, vista por el OI. Se situó el diasporámetro medidor ante el OI y se partió de 12Δ base interna (ver Figura 17). Para realizar la medida se disminuyó lentamente (unas $2\Delta/s$) la potencia prismática base interna (o se aumentó la de base externa). Para evitar incomodidad al sujeto se ajustó la potencia prismática adecuada tanto en el

diasporámetro disociador, como en el medidor antes de situarlos ante sus ojos. Para minimizar los cambios acomodativos (y los cambios de vergencia asociados) se pidió al sujeto que fijara su mirada en la columna inferior de letras, vista por el ojo con el prisma disociador y que se esforzara en mantener nítidos los estímulos en todo momento. El sujeto debía avisar al examinador cuando la columna superior de letras se situara justo sobre la inferior. Se anotó la potencia prismática y la dirección del prisma en esa situación.

VARILLA DE MADDOX VP Y VL EN DIRECCIÓN HORIZONTAL

Como estímulo de fijación se utilizó una luz puntual presentada a 40 cm para cerca y a 6 m para lejos. Durante las medidas se atenuó la luz de sala y se apagó la luz de columna. Se situó en el foróptero la graduación del sujeto y se colocó una varilla de Maddox roja con el eje horizontal ante el OD (ver Figura 18).

Se pidió al sujeto que dirigiera su mirada hacia la luz puntual y si en esta situación no veía superpuestos el punto blanco (visto por el OI) y la línea roja vertical (vista por el OD) se situaba ante el OI el diasporámetro con 12Δ base interna para crear una separación clara entre línea y punto. A continuación se procedía a reducir lentamente (unas $2\Delta/s$) la potencia prismática base interna (o a aumentar la de base externa) hasta que el sujeto indicara que la línea roja vertical y el punto estaban superpuestos. Se anotó la potencia prismática y la dirección de la desviación. Si el sujeto inicialmente apreciaba la luz y la línea superpuestos se anotó directamente “orto”.



Figura 18. Medida de forias con varilla de Maddox en visión próxima utilizando el foróptero.

MÉTODO DE THORINGTON MODIFICADO VP EN DIRECCIÓN HORIZONTAL

Como estímulo de fijación se utilizó la tarjeta de Bernell de medida del desequilibrio muscular en cerca (*Bernell Muscle Imbalance Measure (MIM) card* (Bernell Corp., South Bend, Indiana)) (ver Figura 20). La tarjeta se colocó en la barra del foróptero a 40 cm de modo que los números tuvieran una separación correspondiente a 1Δ . En el centro de la escala graduada hay un agujero a través del cual se proyectó la luz de una linterna puntual dirigida hacia la cara del sujeto (ver Figura 19).

Durante las medidas se mantuvo en un nivel medio la luz de sala y se apagó la luz de columna. Se situó en el foróptero la compensación del sujeto y ante el OD se colocó una varilla de Maddox roja con el eje horizontal.

Se pidió al sujeto que, mientras se esforzaba en mantener las letras nítidas, dirigiera su mirada hacia la luz situada en el centro de la tarjeta y que nos indicara si la línea roja vertical, vista por el OD, estaba sobre el punto luminoso (ortoforia) o estaba a la derecha (endo) o izquierda (exo) respecto al punto de luz. En los dos últimos casos el sujeto nos indicó el número por el que pasaba la línea roja. Se anotó la magnitud (según el paso de 1Δ más próximo) y la dirección de la desviación.



Figura 19. Medida de forias en visión próxima con el método de Thorington modificado.

MEDIDA DE LA DESVIACIÓN VERTICAL

El objetivo fue medir la desviación vertical de los ojos en condiciones de disociación.

VON GRAEFE EN VP EN DIRECCIÓN VERTICAL

Como estímulo de fijación se utilizó una tarjeta colocada a 40 cm en la barra de cerca del foróptero con una única fila de letras correspondiente a un nivel de AV de 0,8. Durante las medidas se mantuvo encendida la luz de sala y la de columna. Se situó en el foróptero la graduación del sujeto y se ajustó adecuadamente la distancia interpupilar.

Como prisma disociador se colocó en el diasporámetro del OD 12Δ base interna, de modo que la fila de letras se desdoblara horizontalmente en dos: la de la derecha, vista por el OD y la de la izquierda, vista por el OI. El diasporámetro medidor se situó ante el OI y se partió de 6Δ base superior. Para realizar la medida se disminuyó lentamente (unas $2\Delta/s$) la potencia prismática base superior (o se aumentó en base inferior). Para evitar incomodidad al sujeto se ajustó la potencia prismática de partida de los diasporámetros antes de situarlos ante sus ojos. Se pidió al sujeto que fijara su mirada en la fila de letras de la derecha, que se esforzara en mantener los estímulos

nítidos en todo momento y que nos avisará cuando la fila de la izquierda se situaba justo a la misma altura que la de la derecha. Se anotó la potencia prismática y la dirección del prisma en esa situación.

VARILLA DE MADDOX VP EN DIRECCIÓN VERTICAL

Como estímulo de fijación se utilizó una luz puntual presentada a 40 cm. Durante las medidas se atenuó la luz de sala y se apagó la luz de columna. Se situó en el foróptero la compensación del sujeto y ante el OD se colocó una varilla de Maddox roja con el eje vertical. Se pidió al sujeto que dirigiera su mirada hacia la luz puntual y si en esta situación no veía superpuestos el punto blanco (visto por el OI) y la línea roja horizontal (vista por el OD) se situaba ante el OI el diasporámetro con 6Δ base superior para crear una separación clara entre línea y punto. A continuación se procedía a reducir lentamente (unas $2\Delta/s$) la potencia prismática base superior (o a aumentar la de base inferior) del diasporámetro hasta que el sujeto nos indicara que la línea roja horizontal y el punto estaban superpuestos. Se anotó la potencia prismática y la dirección de la desviación.

MÉTODO DE THORINGTON MODIFICADO VP EN DIRECCIÓN VERTICAL

Como estímulo de fijación se utilizó la tarjeta de Bernell de medida del desequilibrio muscular en cerca (*Bernell Muscle Imbalance Measure (MIM) card* (Bernell Corp., South Bend, Indiana)) (ver Figura 20). La tarjeta se colocó en la barra del foróptero a 40 cm de modo que los números tuvieran una separación correspondiente a 1Δ . En el centro de la escala graduada hay un agujero a través del cual se proyectó la luz de una linterna puntual hacia la cara del sujeto.

Durante las medidas se mantuvo en un nivel medio la luz de sala y se apagó la luz de columna. Se situó en el foróptero la graduación del sujeto y ante el OD se colocó una varilla de Maddox roja con el eje vertical.

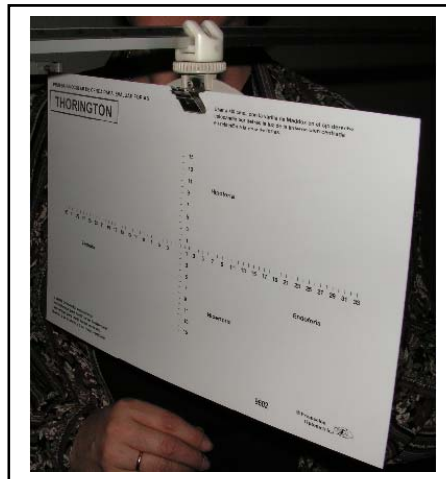


Figura 20. Tarjeta Bernell Muscle Imbalance Measure (MIM) utilizada para medir la foria en visión próxima.

Se pidió al sujeto que dirigiera su mirada hacia la luz situada en el centro de la tarjeta y que nos indicara si la línea roja horizontal vista por el OD estaba sobre el punto luminoso (ortoforia) o estaba a encima (hipo-desviación de OD) o debajo (hiper-desviación de OD) del punto. En los dos últimos casos el sujeto nos indicó el número por el que pasaba la línea roja. Se anotó la magnitud (según el paso de 1Δ más próximo) y la dirección de la desviación.

1.13. Resultados

REPETIBILIDAD

La Tabla 15 muestra los resultados de repetibilidad de las diferentes técnicas de medida de desviación ocular horizontal evaluadas. Las diferencias medias (DM) con signos fueron menores o iguales a $0,7\Delta$ en todas las técnicas estudiadas tanto en VL

como en VP. Las DMs de valores absolutos fueron todas menores de $3,4\Delta$. Ninguna diferencia fue estadísticamente significativa.

Respecto a los coeficientes de repetibilidad (CDR), cabe destacar que en todas las técnicas el CDR de VL fue menor que su CDR de VP; es decir, todas las técnicas presentaron mejor repetibilidad al medir la desviación en VL, que al medir en VP. Los coeficientes de repetibilidad (CDR) para las diferencias en valor absoluto van de $\pm 1,12\Delta$ a $\pm 3,28\Delta$ en VL y de $\pm 2,63\Delta$ a $\pm 6,60\Delta$ en VP. Todos los CDRs aumentaron algo al considerar las diferencias con signos especialmente en las técnicas subjetivas de cerca, llegando hasta intervalos de $\pm 9,26\Delta$ en la técnica de von Graefe en VP.

Tanto en lejos como en cerca, la mejor repetibilidad la presentó el prisma cover test hasta el punto de neutralización del movimiento pues presentó la menor DM y el menor CDR. Si sólo se consideran las técnicas subjetivas, la prueba que presentó mejor repetibilidad en VP, basándonos en su menor CDR, fue la técnica de Thorington modificada.

La Figura 21 muestra la tendencia de valores medios de las distintas técnicas de medida de desviación horizontal tanto en VL como en VP. En lejos, la técnica que como promedio tendió a dar valores más endo fue el método de la varilla de Maddox. En visión próxima, se obtuvieron los valores más endo con la técnica de Thorington modificada. Como es normal, en ambas distancias el punto central entre la neutralización y la inversión es el método que muestra en promedio los valores más exo.

Tabla 15. Repetibilidad en la medida de desviación ocular horizontal en visión lejana y en visión próxima

MEDIDA DE LA DESVIACIÓN OCULAR HORIZONTAL		MEDIA (Δ)	DE (Δ)	DIFERENCIAS CON SIGNO		DIFERENCIAS EN VALOR ABSOLUTO	
				DM (F-I) (Δ) p(test t)	CDR (Δ)	DM (F-I) (Δ)	CDR (Δ)
LEJOS	Cover test (Neutralización)	-0,07	1,15	0,10 (p=0,2)	$\pm 1,17$	0,20	$\pm 1,12$
	Cover test (Punto central)	-1,92	2,12	0,12 (p=0,5)	$\pm 2,84$	0,94	$\pm 2,16$
	Von Graefe	0,52	2,69	-0,43 (p=0,1)	$\pm 3,98$	1,44	$\pm 2,90$
	Varilla de Maddox	3,15	3,56	0,53 (p=0,08)	$\pm 4,53$	1,67	$\pm 3,28$
CERCA	Cover test (Neutralización)	-0,37	2,39	-0,21 (p=0,3)	$\pm 3,18$	0,93	$\pm 2,63$
	Cover test (Punto central)	-2,37	3,45	-0,43 (p=0,2)	$\pm 4,97$	1,68	$\pm 3,80$
	Von Graefe	-1,09	3,90	-0,70 (p=0,2)	$\pm 9,26$	3,36	$\pm 6,60$
	Varilla de Maddox	-1,08	4,91	-0,20 (p=0,7)	$\pm 8,27$	3,15	$\pm 5,45$
	Thorington	0,22	3,56	-0,60 (p=0,2)	$\pm 6,52$	2,20	$\pm 5,00$

Clave: DE = desviación estándar. DM = diferencia media, CDR = coeficiente de repetibilidad, Δ = dioptría prismática. F = final, I = inicial

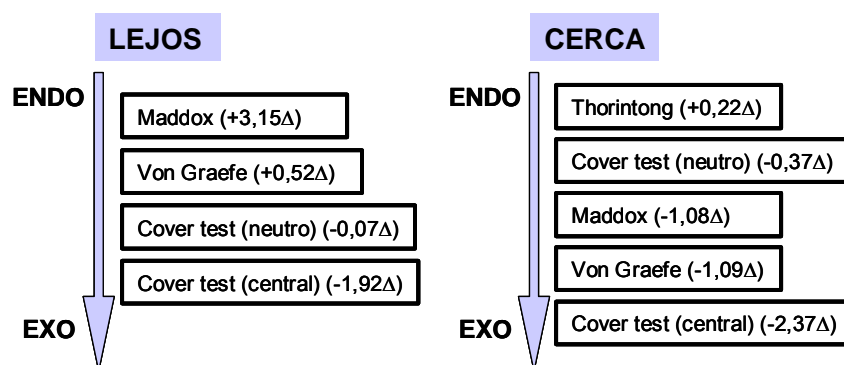


Figura 21. Tendencia de valores medios de las distintas técnicas de medida de desviación horizontal

La Figura 22 muestra las gráficas de Bland y Altman para las diferentes técnicas de medida de la desviación horizontal en lejos y la Figura 23 para las medidas en cerca. El eje-y corresponde a la diferencia con signos entre las medidas de ambas sesiones (Final – Inicial) y el eje-x corresponde a la media de las dos medidas para cada sujeto (en Δ). Por convenio, las exo-desviaciones se muestran a la izquierda con valores negativos y las endo-desviaciones a la derecha con valores positivos. La línea discontinua corresponde a la diferencia media (DM) con signos y las líneas negras continuas corresponden a los límites de acuerdo al 95% ($DM \pm 1,96 \cdot DE$ de las diferencias). Ninguna de las gráficas muestra una tendencia a que la diferencia aumente con el valor prismático, es decir, la repetibilidad de los tests no cambia con la magnitud de la desviación al menos dentro del rango de valores de foria que apareció en la muestra de sujetos estudiada que se extendió aproximadamente de 8Δ de exoforia a 18Δ de endoforia en VL y de 16Δ de exoforia a 16Δ de endoforia en VP. Estas gráficas facilitan la comparación del grado de repetibilidad de las diferentes técnicas, pues cuanto menor sea la diferencia media entre sesiones de medida (DM) y más estrecho sea el intervalo de acuerdo, más repetible será una técnica. Puesto que para cada figura se ha utilizado la misma escala tanto para el eje-x como para el eje-y, es fácil comparar el grado de repetibilidad encontrado para las distintas pruebas. Como se puede ver el intervalo entre los límites de acuerdo al 95% fue más amplio en las técnicas subjetivas, indicando una mayor variación entre sesiones de medida, es decir, una menor repetibilidad.

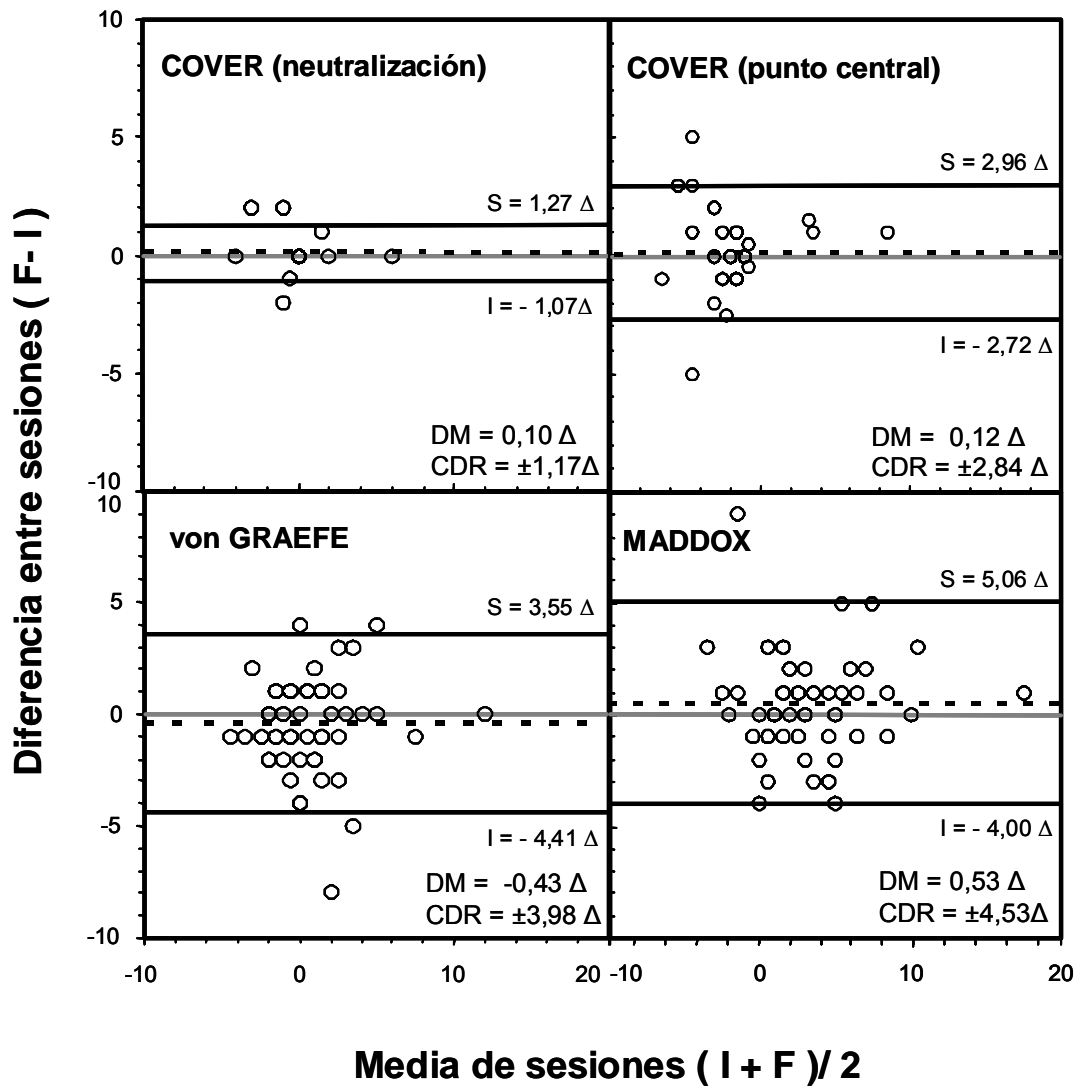


Figura 22. Diferencia media (DM), límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) y coeficiente de repetibilidad (CDR) para la medida de la desviación ocular horizontal en visión lejana.

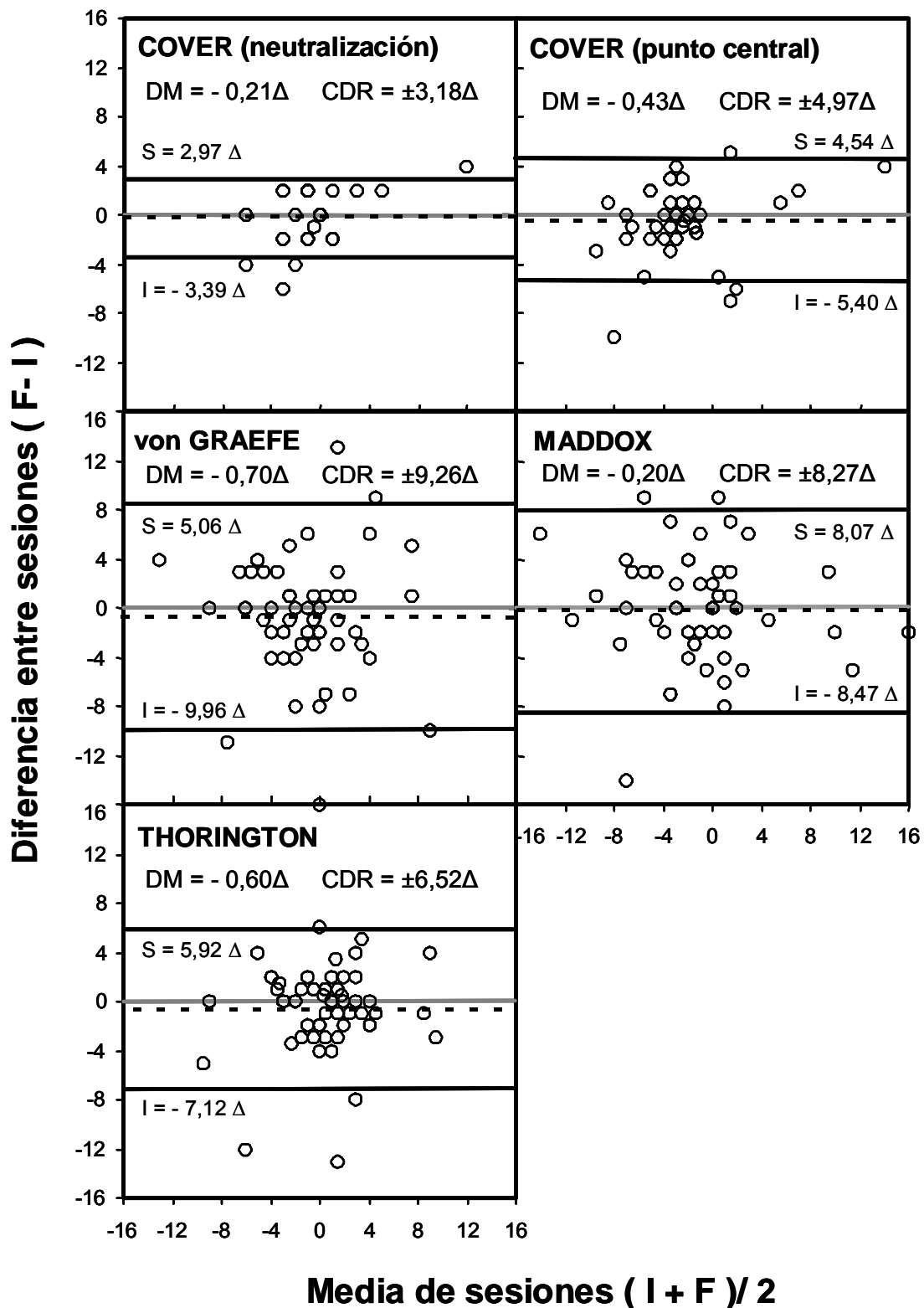


Figura 23. Diferencia media (DM), límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) y coeficiente de repetibilidad (CDR) para la desviación ocular horizontal en visión próxima.

La Tabla 16 muestra los resultados de repetibilidad en la medida de la desviación ocular vertical. Como era de esperar, todas las medidas de desviación vertical estuvieron extremadamente próximas a la ortoforia. Las distribuciones de las diferencias tanto con signos como en valor absoluto no siguieron una distribución normal pues sólo unos pocos sujetos presentaron desviación vertical. Las tres técnicas mostraron una buena repetibilidad con DMs y CDRs clínicamente insignificantes. La Figura 24 muestra las gráficas de Bland y Altman para las diferentes técnicas de medida de la desviación ocular vertical en cerca.

Tabla 16. Repetibilidad en la medida de desviación ocular vertical en cerca

MEDIDA DE LA DESVIACIÓN OCULAR VERTICAL		MEDIA (Δ)	DE (Δ)	DIFERENCIAS CON SIGNO		DIFERENCIAS EN VALOR ABSOLUTO	
				DM (F-I) (Δ) p (test-t)	CDF (Δ)	DM (F-I) (Δ)	CDR (Δ)
CERCA	Von Graefe	0,11	0,28	0,08 (p= 0,2)	$\pm 1,00^*$	0,15	$\pm 1,00^*$
	Varilla de Maddox	0,02	0,20	-0,03 (p = 0,4)	$\pm 0,70^*$	0,07	$\pm 0,70^*$
	Thorington	0,06	0,27	0,00 (p = 1)	$\pm 0,50^*$	0,05	$\pm 0,50^*$

Clave: DE = desviación estándar. DM = diferencia media, CDR = coeficiente de repetibilidad, Δ = dioptría prismática. F = final, I = inicial
 * En los casos donde CDR aparece con un asterisco, se ha sustituido éste por el valor del percentil 95° del valor absoluto de la diferencia

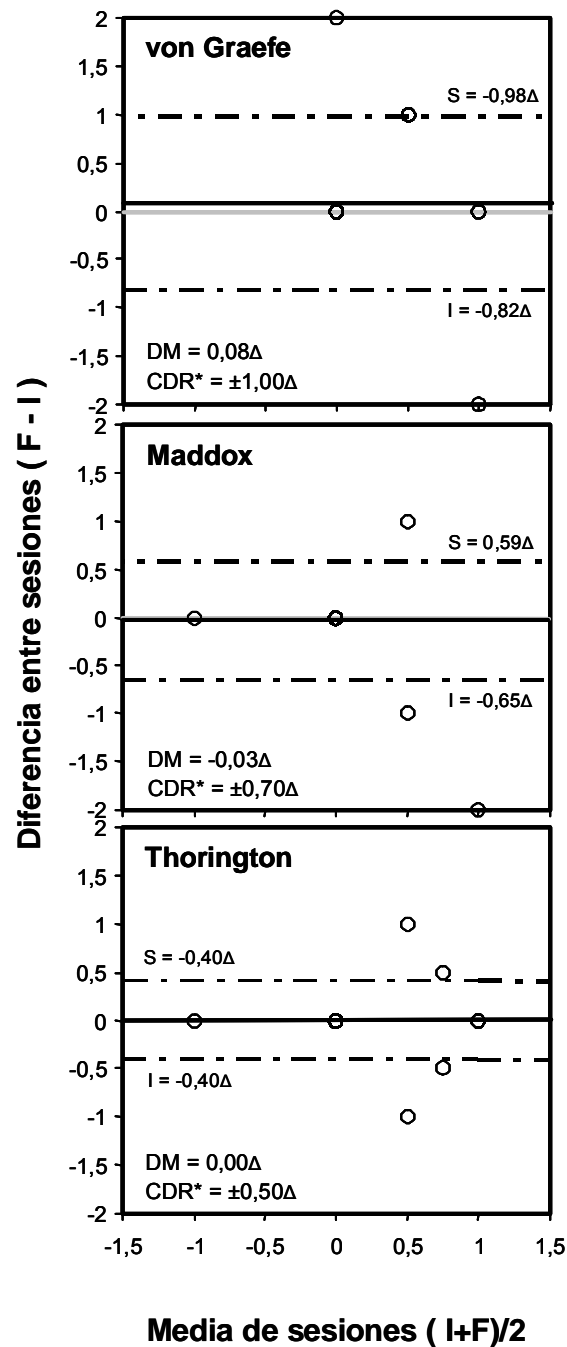


Figura 24. Diferencia media (DM), límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) y coeficiente de repetibilidad (CDR) para la desviación ocular vertical en visión próxima. En los tres métodos la distribución de las diferencias no fue normal y se tomó como coeficiente de repetibilidad el percentil 95 del valor absoluto de las diferencias (CDR*).

CONCORDANCIA

La Tabla 17 muestra los resultados del estudio de concordancia entre las distintas técnicas de medida de la desviación ocular horizontal. Los coeficientes de concordancia (CDC) para las diferencias en valor absoluto fueron de $\pm 1,8\Delta$ a $\pm 5,4\Delta$ en VL y de $\pm 2,0\Delta$ a $\pm 5,4\Delta$ en VP. Todos los CDCs aumentaron notoriamente al considerar las diferencias con signos, llegando hasta intervalos de $\pm 7,5\Delta$ que apareció al comparar el prisma cover test (primera neutralización) con Maddox en VP.

La representación gráfica del estudio de concordancia entre métodos de medida de la foria horizontal se muestra en la Figura 25 (para 6 m) y en la Figura 26 (para 40 cm). Las líneas horizontales representadas en estas gráficas indican la diferencia media con signos (DM) y los límites de acuerdo entre métodos, $DM \pm 1,96 \cdot DE$ de las diferencias. Puesto que para cada figura la escala del eje-y de las diferentes gráficas es la misma, es fácil comparar el grado de acuerdo obtenido entre pares de pruebas. Como es de esperar, los dos puntos finales del cover test muestran el mayor acuerdo entre técnicas, mientras que los demás pares de técnicas reflejan un intervalo de acuerdo bastante parecido, es decir, no hay una notoria diferencia entre el grado de concordancia de las distintas técnicas.

Tanto en visión lejana como en cerca, las gráficas correspondientes al estudio de concordancia entre alguna técnica objetiva, es decir, cover test punto neutro (CTn) ó cover test punto central (CTc) respecto a una de las técnicas subjetivas: von Graefe (VG), varilla de Maddox (MD) ó Thorington modificado (TH) muestran una correlación inversa entre la media y la diferencia entre métodos, de modo que las diferencias entre métodos aumentan al aumentar la medida de la foria horizontal. Este hecho refleja que la concordancia entre métodos objetivos y subjetivos de medida de la foria horizontal empeora al aumentar el valor de la desviación.

Tabla 17. Concordancia entre pruebas que miden la desviación ocular horizontal

MEDIDA DE LA DESVIACIÓN OCULAR HORIZONTAL		DIFERENCIAS CON SIGNO		DIFERENCIAS EN VALOR ABSOLUTO	
		DM (Δ) p(test t)	CDC (Δ)	DM (Δ)	CDC (Δ)
LEJOS	Cover (Neutralización) – Cover (Punto central)	1,86 (p<0,0001)	±2,49	2,05	±1,83
	Cover (Neutralización) – von Graefe	-0,59 (p=0,04)	±4,35	1,72	±2,95
	Cover (Neutralización)– Maddox	-3,22 (p<0,0001)	±5,81	3,40	±5,39
	Cover (Punto central) – von Graefe	-2,45 (p<0,0001)	±3,91	2,54	±3,68
	Cover (Punto central) – Maddox	-5,08 (p<0,0001)	±5,28	5,08	±5,28
	Von Graefe - Maddox	-2,60 (p<0,0001)	±4,66	2,70	±4,66
CERCA	Cover (Neutralización) – Cover (Punto central)	2,00 (p<0,0001)	±2,97	2,27	±2,04
	Cover (Neutralización) – von Graefe	0,72 (p=0,1)	±6,54	2,54	±4,43
	Cover (Neutralización)– Maddox	0,71 (p=0,2)	±7,50	2,73	±5,40
	Cover (Neutralización) - Thorington	-0,59 (p=0,1)	±5,82	2,35	±3,70
	Cover (Punto central) – von Graefe	-1,27 (p=0,005)	±6,69	2,84	±4,42
	Cover (Punto central) – Maddox	-1,28 (p=0,008)	±7,19	3,04	±4,71
	Cover (Punto central) - Thorington	-2,58 (p<0,0001)	±6,13	3,41	±4,26
	Von Graefe - Maddox	-0,01 (p=0,9)	±6,41	2,34	±4,44
	Von Graefe - Thorington	-1,31 (p=0,0002)	±4,95	2,10	±3,73
	Maddox - Thorington	-1,30 (p=0,0008)	±5,64	2,29	±4,23

Clave: DM = diferencia media, CDC = coeficiente de concordancia, Δ = dioptría prismática.

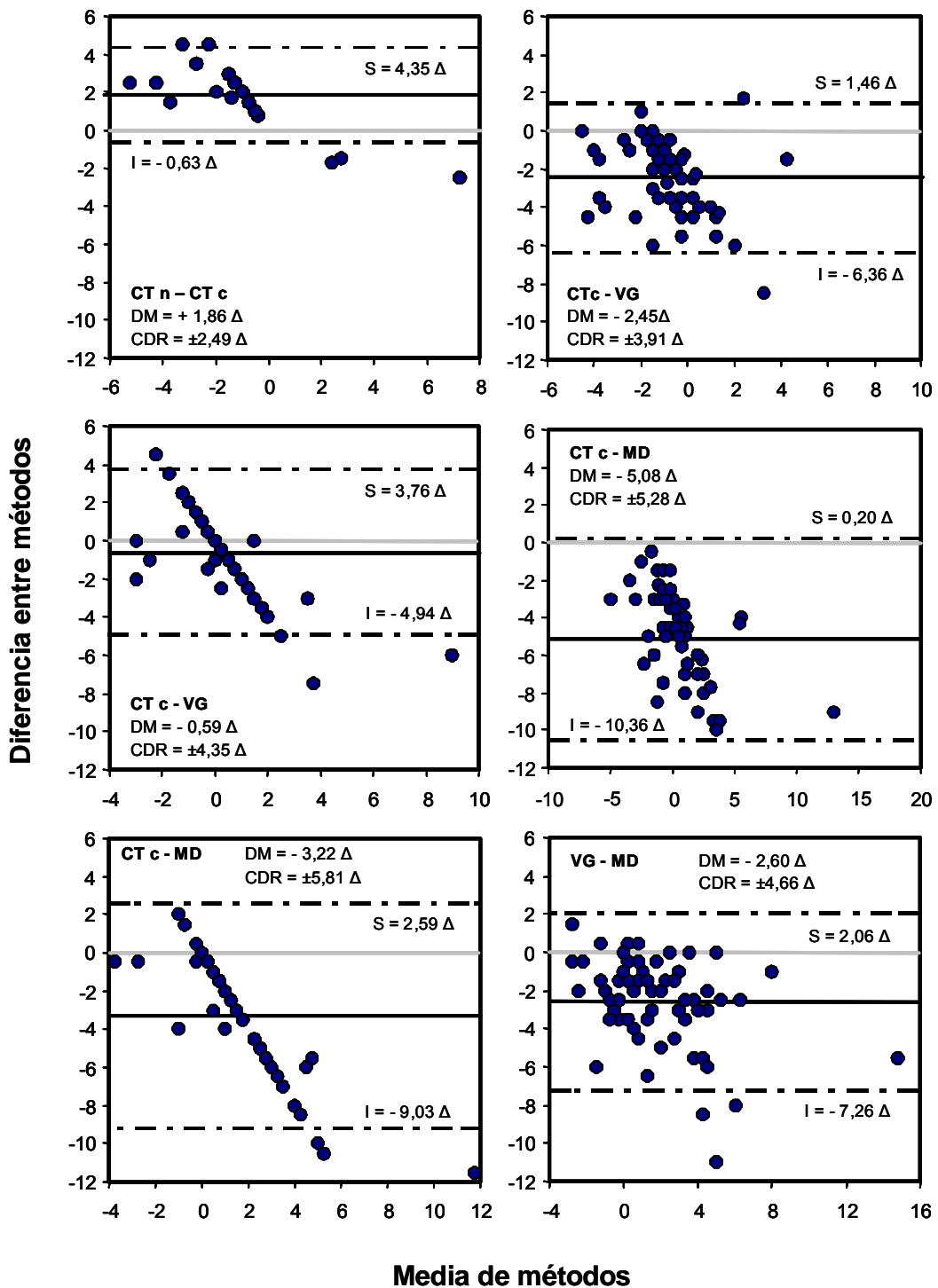


Figura 25. Gráficas de Bland y Altman para el estudio de la concordancia entre métodos que miden la foria horizontal en lejos (6 m). La línea continua representa la diferencia media (DM) entre las medidas obtenidas con los distintos métodos. Las líneas discontinuas indican los límites superior (S) e inferior (I) del intervalo de acuerdo al 95% ($DM \pm 1,96 \cdot DE$).

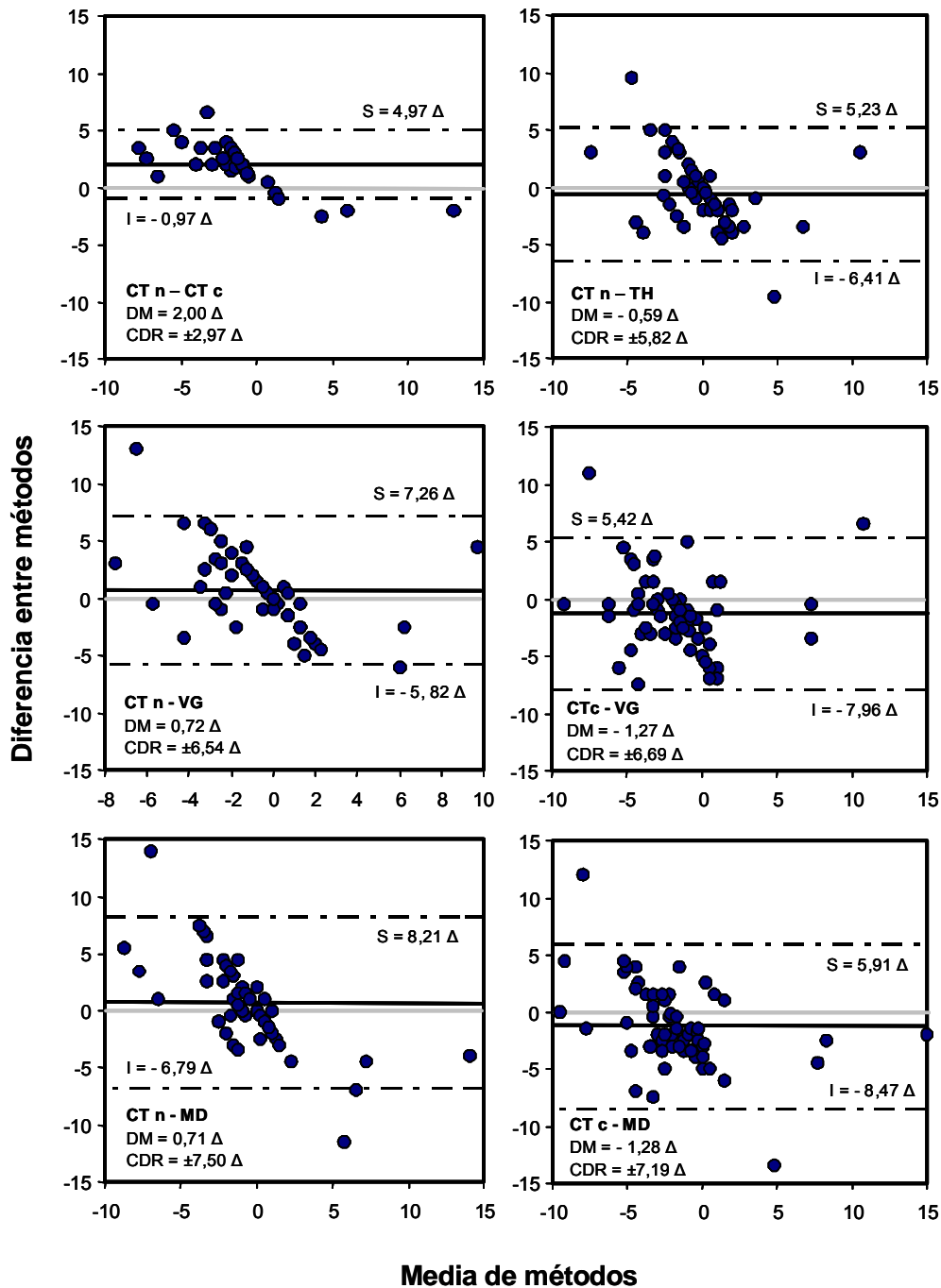


Figura 26. Gráficas de Bland y Altman para el estudio de la concordancia entre métodos que miden la foria horizontal en cerca (40 cm). La línea continua representa la diferencia media (DM) entre los distintos métodos. Las líneas discontinuas indican los límites superior (S) e inferior (I) del intervalo de acuerdo al 95% ($DM \pm 1,96 \cdot DE$).

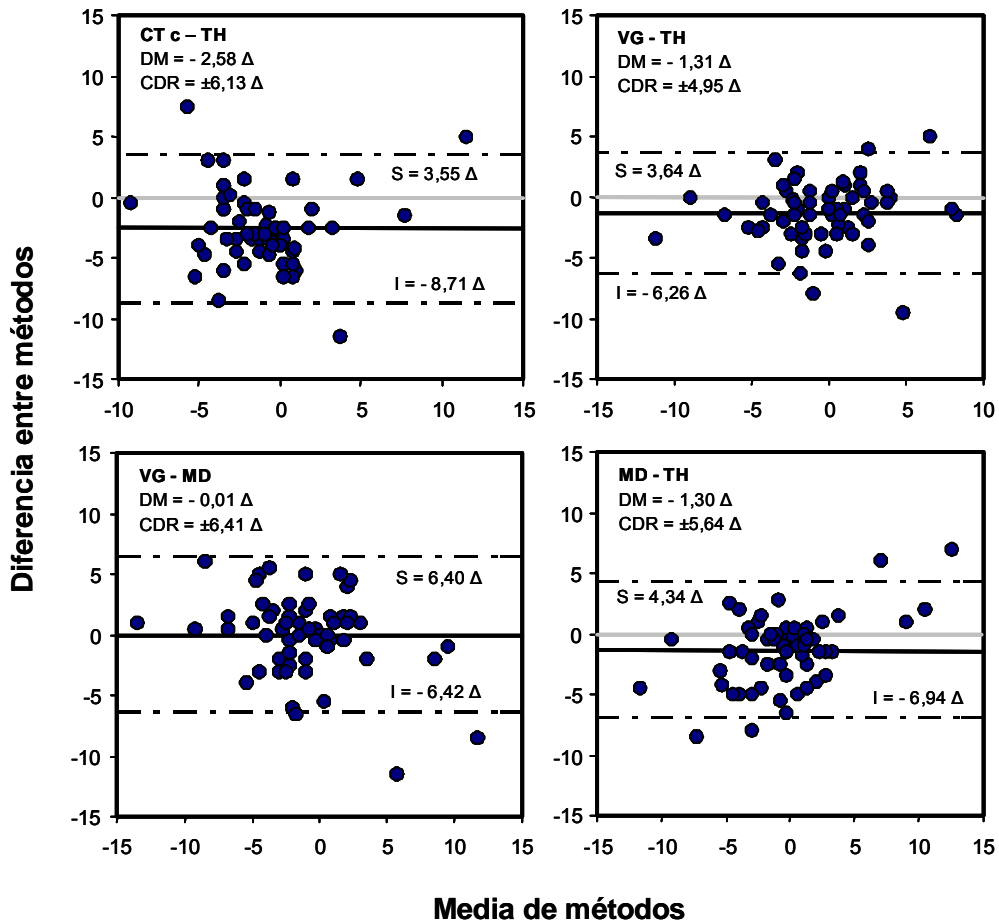


Figura 26. (Continuación).

La Tabla 18 muestra los resultados del estudio de concordancia entre las distintas técnicas de medida de la desviación ocular vertical. Como se puede observar en los tres casos tanto las diferencias medias (DM) como los coeficientes de concordancia (CDC) fueron clínicamente insignificantes (menores o iguales a 1Δ). La representación gráfica del estudio de concordancia entre métodos de medida de la desviación ocular vertical en cerca se muestra en la Figura 27.

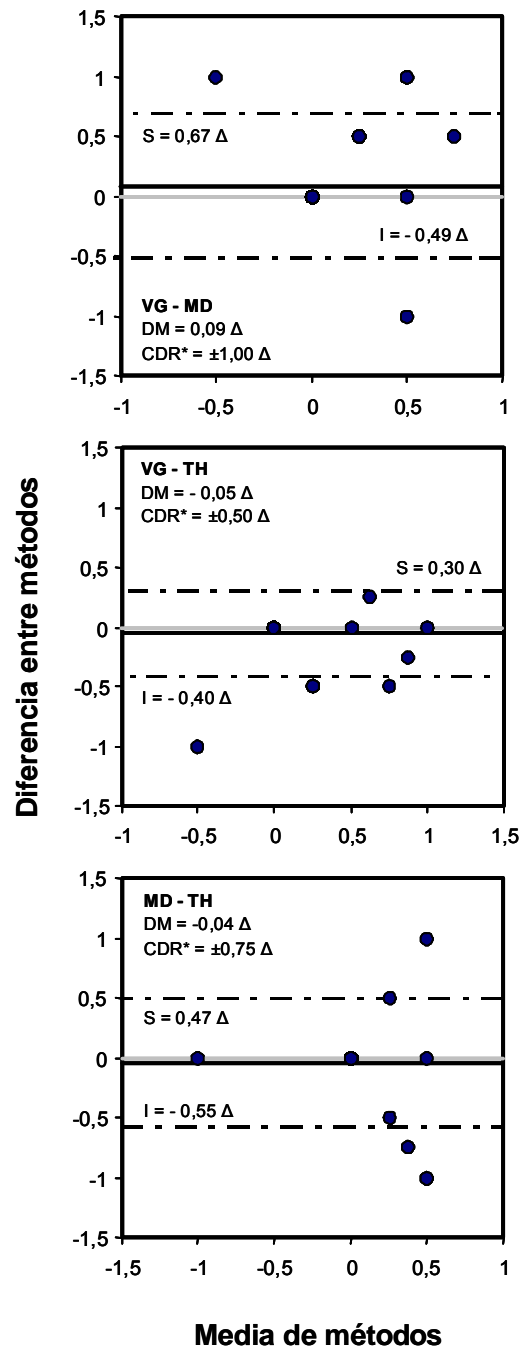


Figura 27. Gráficas de Bland y Altman para el estudio de la concordancia entre métodos que miden la foria vertical en visión próxima (40 cm). La línea continua representa la diferencia media (DM) entre métodos. Las líneas discontinuas indican los límites superior (S) e inferior (I) del intervalo de acuerdo al 95% (DM $\pm 1,96$ DE). En todos los casos la distribución de las diferencias no fue normal y se tomó como coeficiente de concordancia el percentil 95 del valor absoluto de las diferencias (CDC*).

Tabla 18. Concordancia entre pruebas que miden la desviación ocular vertical

MEDIDA DE LA DESVIACIÓN OCULAR VERTICAL		DIFERENCIAS CON SIGNO		DIFERENCIAS EN VALOR ABSOLUTO	
		DM (Δ) p(test-t)	CDC (Δ)	DM (Δ)	CDC (Δ)
CERCA	Von Graefe - Maddox	0,09 (p = 0,03)	$\pm 1,00^*$	0,12	$\pm 1,00^*$
	Von Graefe – Thorington	-0,05 (p = 0,04)	$\pm 0,50^*$	0,06	$\pm 0,50^*$
	Maddox - Thorington	-0,04 (p = 0,2)	$\pm 0,75^*$	0,09	$\pm 0,75^*$

Clave: DM = diferencia media, CDC = coeficiente de concordancia, Δ = dioptría prismática.
 * En los casos donde CDC aparece con un asterisco, se ha sustituido éste por el valor del percentil 95º del valor absoluto de la diferencia

ANOVA

Los resultados de un ANOVA de medidas repetidas con dos factores (método y sesión) para las medidas de desviación horizontal con signos mostraron diferencias significativas en función del método aplicado tanto en VL ($F_{3,180}=98,27$; $p<0,0001$) como en VP ($F_{4,240}=11,741$; $p<0,0001$). Sin embargo, no aparecieron diferencias significativas entre las dos sesiones de medida ni en VL ($F_{1,60}=0,43$; $p=0,5$), ni en VP ($F_{1,60}=2,56$; $p=0,1$). Tampoco se encontraron efectos de interacción entre ambos factores ni en VL ($F_{3,180}=3,72$; $p=0,06$), ni en VP ($F_{4,240}=0,33$; $p=0,8$). El análisis *post-hoc* utilizando la técnica de Scheffé indicó una diferencia significativa entre las medias de cada uno de los métodos de VL excepto al comparar el cover test (punto neutro) con la técnica de von Graefe. Sin embargo, en VP sólo se mostraron significativas las diferencias entre los dos valores del cover test, punto neutro y punto central, y del cover test (punto central) con el test de Thorington modificado.

Los resultados de un análisis ANOVA equivalente al anterior para las medidas de desviación vertical en VP no mostraron diferencias significativas en función del método aplicado ($F_{2,120}=3,65$; $p=0,06$) ni entre las dos sesiones de medida ($F_{1,60}=0,26$; $p=0,6$). Tampoco hubo efectos de interacción ($F_{2,120}=2,67$; $p=0,08$).

1.14. Discusión

La medida de la desviación ocular es posiblemente uno de los aspectos optométricos más estudiados en los últimos 50 años. En un intento de aportar claridad, tras seleccionar y analizar los estudios más relevantes sobre el tema, se han elaborado dos tablas, la Tabla 19 que recopila los principales estudios de repetibilidad de la medida de forias y la Tabla 20 que resume los estudios sobre comparación de pruebas. Los resultados de varios de los estudios que aparecen en estas tablas se discutirán más adelante comparándolos con los obtenidos en esta investigación. Como se puede observar en estas tablas recopilatorias es difícil extraer ideas unánimes, pero sí parece que tienden a repetirse las siguientes conclusiones generales respecto a las distintas pruebas de medida de desviación ocular analizadas en el estudio presente:

- La medida de desviación horizontal con varilla de Maddox no es una prueba muy recomendable debido a las condiciones artificiales bajo las que se aplica y porque una luz puntual no es un buen estímulo acomodativo.
- Entre las pruebas subjetivas, el test de Thorington modificado se ha mostrado en sucesivos estudios como el método más repetible para medir forias.
- La mayor repetibilidad de las medidas en visión lejana podría asociarse a la menor variabilidad de la respuesta acomodativa en lejos que en cerca.
- Muchos de los estudios previos de repetibilidad utilizan análisis de correlación, que indica asociación, más que concordancia entre los distintos resultados.
- La medida de forias en el espacio libre (gafa de pruebas) tiende a mostrarse más repetible que la medida equivalente con foróptero.
- La tendencia del método de von Graefe a dar más exo podría deberse a la adaptación de vergencias asociada a usar 12Δ BI como punto de partida del diasporámetro medidor.

Tabla 19. Resumen de estudios de repetibilidad en la medida de forias

ESTUDIO	PRUEBAS EVALUADAS ^a	DIFERENCIAS METODOLÓGICAS ^b	ANÁLISIS ESTADÍSTICO	CONCLUSIONES PRINCIPALES
Scobee y Green ^{95 c}	Maddox horizontal y vertical en VL y VP	VP a 33 cm Medidas tomadas por 2 examinadores en días distintos (4 medidas por día) N=50	Análisis de correlación Maddox horiz.: r = 0,95 Maddox vert.: r = 0,92	Además del estudio de repetibilidad interexaminador también estudiaron el efecto de varias variaciones metodológicas y vieron que no había diferencias clinicamente significativas entre: - Medir tapando y destapando el ojo con el diaporámometro medidor o no (N=100) - Medir con luz ambiente o en la oscuridad (N=60) - Colocar la varilla de Maddox ante el ojo dominante o ante el no dominante - Usar una varilla de Maddox roja o blanca
Hirsch y Bing ^{96 c}	Von Graefe horizontal VP Thorington modificado horizontal VP	Von Graefe con E Snellen grande o con fila de letras correspondiente al mínimo legible Maddox blanco N=38 conocedores de las pruebas	Bland y Altman en valor absoluto ^d Von Graefe (letras): Ex 1: 2,2±3,53Δ Ex 2: 2,0±3,53Δ Thorington modificado: Ex 1: 1,4±3,53Δ Ex 2: 1,2±2,16Δ	- Thorington es el test más repetible - En von Graefe, no hay diferencias significativas entre los distintos test de fijación o entre medidas repetidas
Doherty y Geraci ^{97 c}	von Graefe horizontal VP	Fijando uno u otro ojo la tarjeta N=40	Diferencia media en valor absoluto: 0,7Δ	- Para sujetos con diferencia de equivalente esférico entre ambos ojos no superior a 0,12 D no hay diferencias clinicamente significativas entre fijar uno u otro ojo
Morgan ^{98 c}	Maddox horizontal VL Von Graefe horizontal VP	Tapando y destapando el ojo con el diaporámometro medidor Toman 5 medidas por cada test, calculan la DE de cada sujeto y extraen la DE media	Diferencia media Maddox VL: 1,0Δ Von Graefe VP: 2,1Δ	- Von Graefe muestra menor repetibilidad que varilla de Maddox

Tabla 19. Continuación

ESTUDIO	PRUEBAS EVALUADAS ^a	DIFERENCIAS METODOLÓGICAS ^b	ANÁLISIS ESTADÍSTICO	CONCLUSIONES PRINCIPALES
Morris ^{90 c}	Maddox horizontal VP Von Graefe horizontal VP Thorington horizontal VP	VP a 35,5 cm Maddox tapando intermitente un ojo Cada test se aplica 4 veces en días distintos y siempre en el mismo orden Con (1ª y 2ª sesión) o sin (3ª y 4ª sesión) efecto quinesístico (tocar la tarjeta) N=20 no conocedores de las pruebas	Bland y Altman Maddox: $1,6 \pm 3,6\Delta$ Von Graefe: $1,9 \pm 2,9\Delta$ Thorington: $1,1 \pm 1,6\Delta$	- Thorington modificado es más repetible que von Graefe o Maddox. - Maddox es la prueba más afectada por el uso de apoyo táctil (un poco más endo $\approx 1\Delta$). Aunque las diferencias no son clínicamente significativas
Rouse <i>et al.</i> ⁷	Von Graefe horizontal VP	Tapando y destapando el ojo con el diaporámetro medidor N=20 (10-11 años)	Bland y Altman ^d Ex 1: $\pm 6,78\Delta$ Ex 2: $\pm 7,64\Delta$	- Von Graefe en cerca tiene una repetibilidad baja
Johns <i>et al.</i> ⁹¹	Cover test VP primera neutralización Cover test VP media entre primera neutralización e inversión N=72 (19-58 años)	N = 72 (19-58 años)	Bland y Altman ^d Con signos: DM < $0,5\Delta$ CDR < $\pm 5,0\Delta$ Valor absoluto: DM < $1,9\Delta$ CDR < $\pm 3,4\Delta$	- El cover test es una prueba muy repetible en todas sus variantes

Tabla 19. Continuación

ESTUDIO	PRUEBAS EVALUADAS ^a	DIFERENCIAS METODOLÓGICAS ^b	ANÁLISIS ESTADÍSTICO	CONCLUSIONES PRINCIPALES
Lam <i>et al.</i> ¹⁰⁰	Thorington modificado horizontal VP Maddox horizontal VP	Miden en Thorington con y sin prisma Comparan medidas en gafa de prueba y foróptero N=40 (19-22 años) Sólo sujetos con la misma dirección de foria en los distintos métodos(N=31) Usan en el análisis la media de tres medidas	Porcentaje de variación Thorington: 69,5% Maddox: 26,5%	- Mayor variabilidad en el test de Thorington modificado que en Maddox - El uso de prismas para medir la desviación frente al uso de la escala graduada en el test de Thorington modificado no afecta al resultado de la foria medida.
Casillas y Rosenfield ¹⁰¹	Von Graefe Maddox Thorington modificado (horizontal y vertical, VP y VL) N=60 (20-34 años)	Comparan medidas en gafa de prueba y foróptero Cada sesión empezó con 5 min de oclusión de un ojo Maddox: miden partiendo de cero	Bland y Altman ^d . Foróptero. Horizontal VG: VL: ±4,16; VP: ±5,28 Δ MD: VL: ±3,91, VP: ±5,73 Δ TH: VP: ±3,28 Δ Foróptero. Vertical VG: VP: ±0,51 Δ MD: VP: ±0,35 Δ TH: VP: ±0,25 Δ ANOVA: diferencia significativa foróptero vs. gafa de pruebas, pero no entre sesiones	- Las desviaciones verticales medias estuvieron muy próximas a la ortoforia. - Todas las técnicas mostraron mejor repetibilidad con gafa de pruebas que con foróptero - Von Graefe fue la técnica de medida de desviación vertical y horizontal, VL y VP, menos repetible. - Maddox con gafa de pruebas es el procedimiento más repetible para medir desviaciones horizontales en cerca

^a Sólo se hace referencia explícita a aquellas pruebas que también se han analizado en el estudio presente
^b Se incluyen algunas diferencias metodológicas importantes en la aplicación de las pruebas que podrían influir en las conclusiones de estos estudios respecto a los resultados de nuestra investigación
^c Datos tomados del artículo de revisión de Schroeder *et al.* ⁹²
^d Ex 1 = examinador 1. Ex 2: examinador 2. DM = Diferencia media. CDR = Coeficiente de repetibilidad. VG = von Graefe, MD = Maddox, TH = Thorington.
 VL = visión lejana, VP = visión próxima

Tabla 20. Resumen de estudios que comparan las pruebas de medida de forias.

ESTUDIO	PRUEBAS EVALUADAS ^a	DIFERENCIAS METODOLÓGICAS ^b	ANÁLISIS ESTADÍSTICO	CONCLUSIONES PRINCIPALES
Scobee y Green ^{95 c}	Cover test subjetivo horizontal VL y VP Maddox horizontal en VL y VP Von Graefe horizontal VL y VP Thorington modificado horizontal VP	En el cover test el sujeto informaba del movimiento relativo del objeto al compensar (CT subjetivo) Maddox tapando intermitente un ojo VP a 33 cm N=50	Análisis de correlación ^d : VL: VG – Cover: r = 0,84 MD – Cover: r = 0,81 MD – VG: r = 0,85 VP: VG – Cover: r = 0,72 MD – Cover: r = 0,48 TH – Cover: r = 0,77 MD – VG: r = 0,41 MD – TH: r = 0,35 TH – VG: r = 0,88	En visión lejana: - Las tres pruebas dan aproximadamente la misma foria promedio y presentan la misma dispersión en torno a la media En visión próxima: - El cover test subjetivo y Maddox dan mayor cantidad de exo desviación que von Graefe y Thorington. - Basándose en las medias y las desviaciones estándar las cuatro pruebas no son equivalentes
Hirsch y Bing ^{96 c}	Von Graefe horizontal VP Thorington modificado horizontal VP	Von Graefe con E Snellen grande o con fila de letras correspondiente al mínimo legible N=38 conocedores de las pruebas	Análisis de correlación: Von Graefe (fila letras) vs. Thorington: r = 0,92	- Thorington da 1,3Δ menos exo que von Graefe con letras - En von Graefe, no hay diferencias significativas entre los distintos test de fijación
Hirsch ^{102 c}	Von Graefe horizontal VP Thorington modificado horizontal VP	Von Graefe con 8Δ Base Inferior en OI y 12 Δ Base Interna en OD	Análisis de correlación: Optometrista experimentado: r = 0,84 Estudiante: r = 0,74	- No hay diferencias significativas entre ambas pruebas (Thorington da 1Δ menos exo que von Graefe), pero Hirsch recomienda utilizar el test de Thorington por su mayor simplicidad
Weymouth ^{103 c}	Cover test subjetivo horizontal VL Maddox horizontal en VL Von Graefe horizontal VL	CT: sujeto informaba del movimiento relativo del objeto (CT subjetivo) VL a 7 m Hacen 5 medidas con cada test N=12	Bland y Altman ^e Von Graefe vs. Maddox: -0,2±1,7Δ Von Graefe vs. CT subjetivo: +0,3±2,2Δ Maddox vs. cover test subjetivo: -0,4±1,9Δ	- Todos los métodos ofrecen medidas precisas y consistentes, cuya diferencia es tan pequeña que la elección de un método u otro es indiferente

Tabla 20. Continuación

ESTUDIO	PRUEBAS EVALUADAS ^a	DIFERENCIAS METODOLÓGICAS ^b	ANÁLISIS ESTADÍSTICO	CONCLUSIONES PRINCIPALES
Soderberg ^{104, c}	Maddox horizontal y vertical VL Von Graefe horizontal y vertical VL	Muestra de pacientes clínicos (N=100)	Análisis de correlación Horizontal: r = 0,97 Vertical: r = 0,91	- Ambos métodos son tan parecidos que uno podría ser sustituido por el otro
Daum ¹⁰⁵	Cover test horizontal primera neutralización Maddox horizontal Von Graefe horizontal	Miden a 1m (demanda de acomodación=1D) 2 grupos (21-40 años) N ₁ =343 visión normal N ₂ =18 estrábicos	Análisis de correlación ^d CT – MD: Gr1: r = 0,89 Gr2: r = 0,94 CT – MD: Gr1: r = 0,89 Gr2: r = 0,95 CT – MD: Gr1: r = 0,92 Gr2: r = 0,99	- Las variaciones entre métodos fueron mayores en sujetos con visión binocular normal que en estrábicos - La correlación entre pruebas subjetivas fue mayor que entre la prueba objetiva y las subjetivas
Calvin <i>et al.</i> ¹⁰⁶	Cover test estimación horizontal VL y VP Von Graefe horizontal en VL y VP	2 examinadores estudiantes universitarios de 4º curso N=25 pacientes clínicos (21-64 años) N=28 estudiantes universitarios de 1º y 2º curso	Análisis de correlación VL: r = 0,77 VP = 0,71 Bland y Altman VL: ±6,9Δ VP: ±8,3Δ	- Ambos métodos no pueden considerarse intercambiables excepto en altas forias

Tabla 20. Continuación

ESTUDIO	PRUEBAS EVALUADAS ^a	DIFERENCIAS METODOLÓGICAS ^b	ANÁLISIS ESTADÍSTICO	CONCLUSIONES PRINCIPALES
Johns et al. ⁹¹	Cover test VP primera neutralización Cover test VP media entre primera neutralización e inversión.	N=72 (19-58 años)	Bland y Altman ^d Con signos: DM < 1,0Δ CDC < ±3,6Δ Valor absoluto (p<0,0005) DM < 1,6Δ CDC < ±2,6Δ	- La diferencia media entre ambos puntos finales del cover test es menor de 2Δ tanto en el promedio con signos como en valor absoluto
Lam et al. ¹⁰⁰	Thorington modificado horizontal VP Maddox horizontal VP	Miden en Thorington con y sin prisma Comparan medidas en gafa de prueba y foróptero N=40 (19-22 años) Sólo sujetos con la misma dirección de foria en los distintos métodos (N=31) Usan en el análisis la media de 3 medidas	ANOVA de medidas repetidas Compara valor medio ±DE Thorington (-2,9± 5,2 Δ) vs. Maddox (-2,9±5,7 Δ)	- No hubo diferencias significativas entre los distintos métodos (F=0,73, p>0,05) - No hubo diferencias significativas entre la medida con foróptero y con gafa de pruebas - El uso de prismas para compensar la desviación en Thorington no afecta a los resultados
Casillas y Rosenfield ¹⁰¹	Von Graefe Maddox Thorington modificado (horizontal y vertical VP y VL)	Comparan medidas en gafa de prueba y foróptero Cada sesión empezó con 5 min de oclusión de un ojo Maddox: miden partiendo de cero N=60 (20-34 años)	ANOVA: Horizontal: diferencia significativa entre cada uno de los métodos	- El test de Thorington modificado tiende a dar valores más endo y el de von Graefe los más exo - Vertical: Las diferencias entre métodos fueron clínicamente insignificantes.

^a Sólo se hace referencia explícita a aquellas pruebas que también se han analizado en el estudio presente

^b Se incluyen algunas diferencias metodológicas importantes en la aplicación de las pruebas que podrían influir al extraer conclusiones de estos estudios respecto a los resultados de nuestra investigación

^c Datos tomados del artículo de revisión de Schroeder et al.⁹²

^d Ex 1 = examinador 1. Ex 2: examinador 2. DM = Diferencia media. CDC = Coeficiente de concordancia. CT = cover test. VG = von Graefe, MD = Maddox, TH = Thorington.

VL = visión lejana, VP = visión próxima. Gr1 = grupo 1. Gr2 = grupo 2

^e -: primer test más exo que el segundo. +: primer test más endo que el segundo.

REPETIBILIDAD

Varios de los estudios que evalúan la repetibilidad de las distintas pruebas que miden la desviación ocular se remontan a hace más de 40 años⁹⁵⁻⁹⁹. Más recientemente fue publicado un artículo de revisión⁹² sobre la repetibilidad de las pruebas que miden forias del que se puede extraer la siguiente conclusión: como es de esperar no todas las pruebas muestran la misma repetibilidad y además una misma prueba puede mostrar diferencias significativas en el grado de repetibilidad de un estudio a otro. Schroeder *et al.*⁹² comentan que estas diferencias podrían atribuirse a:

- Sesgos inter-examinador
- Variaciones o errores del examinador durante la aplicación de la técnica
- Mal control acomodativo
- Distracción, fatiga o diferente capacidad de respuesta entre los sujetos analizados.

Otros estudios^{105, 107} recogen más aspectos que también podrían influir en la variabilidad de las medidas de foria por ejemplo:

- Mala calibración de los instrumentos de medida
- Preferencia por los números pares frente a los impares⁹⁸
- Influencia sobre la dirección y magnitud de la foria medida de otras pruebas previamente realizadas
- Variaciones en las condiciones de la prueba como dispersión cromática a través de los prismas o deslumbramiento.

Para todos los métodos estudiados se encontró que las variaciones entre sesiones (CDR) fueron apreciablemente mayores en los valores de la foria de cerca que en la foria de lejos. Una posible causa de estas mayores diferencias podría encontrarse en

el nexo entre los cambios en la acomodación y la convergencia (y por tanto en la foria). La acomodación puede fluctuar incluso en ausencia de estímulo acomodativo¹⁰⁸ y podría suceder lo mismo con la convergencia. Además, es de esperar que estas fluctuaciones sean mayores en cerca que en lejos. Rainey *et al.*⁹³ plantearon que aunque la amplitud del intervalo de acuerdo (CDR) aceptable está sujeto a debate, los intervalos superiores a $\pm 6,7\Delta$ podrían ser considerados como clínicamente inaceptables

Aunque todos los sujetos del estudio presente presentaban visión binocular normal, sin estrabismo o ambliopía, es clínicamente importante que una prueba para medir desviaciones sea capaz de identificar y medir con precisión las forias altas, dada su elevada probabilidad de generar síntomas clínicos como dolor de cabeza y astenopía. El análisis de las gráficas (Figura 22 y Figura 23) parece indicar que los resultados de este estudio no dependen de la magnitud de la foria medida, pues las diferencias no son mayores en aquellos sujetos con forias más altas. No obstante, sería apropiado repetir el estudio con una muestra de sujetos que presentaran problemas de visión binocular.

A continuación se expondrán algunas consideraciones particulares sobre la repetibilidad de las pruebas estudiadas:

Cover test

En el cover test incluso aunque se hubieran encontrado diferencias estadísticamente significativas, sólo se considerarían como diferencias clínicamente significativas las que fueran mayores o iguales a 2Δ . Este nivel de significación se estableció basándonos en diversos estudios que encontraron que el menor movimiento ocular que una persona puede detectar de forma fiable sin ayuda es 2Δ en condiciones

ideales de examen^{109, 110}; aunque, en las condiciones clínicas habituales bajo las que se aplican estas pruebas, el límite puede estar más próximo a $3-4\Delta$ ¹¹¹.

El estudio de repetibilidad para el cover test reflejó que en todas las condiciones las DMs tanto con signos como en valor absoluto en ningún caso llegaron siquiera a 1Δ , de modo que se puede concluir que la técnica del prisma cover test es un método muy repetible de medida de la desviación tanto si se considera la primera neutralización como si se utiliza la media entre neutralización e inversión. Estos resultados están de acuerdo con los encontrados por Johns *et al.*⁹¹ en la aplicación del cover test en cerca.

Por otra parte, Sparks¹¹² vio que la posición del examinador al medir la foria en cerca con el cover test (posición central, justo detrás del estímulo de fijación *versus* desde un lateral) no era relevante en sujetos endofóricos y ortofóricos. Sin embargo, en los sujetos exofóricos, sus resultados sugirieron que cabría esperar mayor cantidad de exoforia si el examinador se situaba lateralmente respecto a la posición del estímulo de fijación (4Δ más de media). Clark *et al.*¹¹³ reprodujeron el estudio de Sparks¹¹² y encontraron la misma tendencia de valores: cambios no significativos para ortofóricos y endofóricos y, en exofóricos, aumento de la exoforia al medir desde 30° a la derecha (aunque su diferencia media fue menor: $1,3\Delta$). De estos resultados se desprende que la posición que adopte el examinador puede influir en la repetibilidad del cover test en cerca y será importante evaluar desde la misma posición a los sujetos. En el estudio presente siempre se midió la foria en cerca estando el examinador centrado.

Comentaremos a continuación algunos aspectos metodológicos que podrían acotar el alcance de las conclusiones extraídas de este estudio, por ejemplo: (1) El examinador que realizó las medidas contaba con gran experiencia clínica que le facilitó la detección de pequeños movimientos oculares y que le llevó a tener especial cuidado en disociar bien a cada sujeto y en no permitir visión binocular durante las medidas.

Ello lleva a pensar que los resultados puede que no fueran los mismos en clínicos menos experimentados o en estudiantes. (2) La repetibilidad intra-examinador, especialmente en el cover test por ser una prueba objetiva, podría verse algo sesgada por la incapacidad de enmascarar totalmente al examinador que podría recordar esporádicamente los resultados de algún sujeto. Sin embargo, el alto número de sujetos examinados, combinado con el periodo de tiempo transcurrido entre sesiones de medidas y, por supuesto, la ocultación de los datos de la primera tanda de medidas al llevar a cabo la segunda tanda; hacen que el hecho de que el examinador pudiera recordar los resultados de medidas previas fuera extremadamente difícil. Por ello, creemos que las deficiencias en el enmascaramiento no han afectado a los resultados de forma apreciable.

Von Graefe

Rouse *et al.*⁷ estudiaron la repetibilidad intra-examinador para la medida de foria en cerca con la técnica de von Graefe y encontraron unos intervalos de acuerdo al 95% con signos de $\pm 6,78\Delta$ para el examinador 1 y de $\pm 7,64\Delta$ para el examinador 2. Es decir, los CDRs fueron bastantes altos y similares a los del estudio presente ($\pm 6,60\Delta$ con valores absoluto y $\pm 9,26\Delta$ con signo).

Rainey *et al.*⁹³ evaluaron la repetibilidad inter-examinador y mostraron que la técnica de von Graefe en cerca era la técnica menos repetible de las aplicadas usualmente, con un CDR clínicamente elevado ($\pm 6,7\Delta$). Estudios posteriores^{7, 90} también estudiaron la repetibilidad inter-examinador y obtuvieron unos CDRs similares a los de Rainey.

A la vista de estos resultados se puede concluir que la repetibilidad inter- e intra-examinador para la medida de foria con la técnica de von Graefe en cerca es baja con intervalos de acuerdo superiores a $\pm 6\Delta$. No está clara la causa por la que los

resultados de esta prueba son tan variables pero entre las razones propuestas está el entorno artificial producido por el foróptero y el tiempo necesario para administrar la prueba que podría permitir cambios en el nivel de convergencia fusional del sujeto y, por tanto, en la foria medida⁹³.

Maddox y Thorington

La medida de la desviación horizontal con varilla de Maddox no parece una prueba muy recomendable, especialmente en cerca. Entre los factores que pueden contribuir a ello están las condiciones artificiales bajo las que se aplica la prueba y el hecho de que una luz puntual no es un buen estímulo acomodativo y hace más difícil mantener el nivel de acomodación necesario para una determinada distancia de observación. Típicamente una luz genera una respuesta acomodativa significativamente menor de la esperada para la distancia del test, lo que puede llevar a hipoestimar una endodesviación o a hiperestimar una exodesviación¹¹⁴. Algunos de estos problemas se han solucionado con el método modificado de Thorington donde la tarjeta de fijación incluye pequeñas letras y números que ayudan a controlar la respuesta acomodativa. No obstante, aun quedan algunos problemas potenciales por solucionar¹⁰⁰, por ejemplo, el uso de varilla de Maddox hace que un ojo perciba una línea roja que es muy diferente a la escala graduada que ve el otro ojo. Esta rivalidad binocular podría influir en la respuesta acomodativa y para paliarlo se debe pedir al paciente que trate de mantener la escala nítida en todo momento. Por otra parte, si se mide la foria de cerca con el foróptero, se impide la mirada inferior lo que podría dar pistas proximales erróneas al sistema acomodativo.

A pesar de los aspectos metodológicos comentados en el párrafo anterior, el test de Thorington modificado se ha mostrado como el más repetible en diversos estudios^{90, 92, 93, 96, 99, 102}, además, es bastante fácil de administrar, requiere un mínimo equipamiento

y los sujetos entienden fácilmente en qué consiste la prueba y son capaces de responder sin dificultades. Nuestros resultados también sitúan al tests de Thorington modificado como la prueba subjetiva más repetible.

Foria vertical

Como se ha comentado anteriormente, la medida de la desviación vertical tanto en VL como en VP presentó unas diferencias medias tanto con signos como en valor absoluto muy próximas a la ortoforia. Al comparar los resultados de las distintas técnicas de medida (ver Tabla 16), se puede ver que la técnica de von Graefe presentó el CDR más extenso, es decir, aunque la diferencia fue clínicamente insignificante, se muestra como la técnica menos repetible. Este resultado está de acuerdo con el encontrado por Casillas y Rosenfield¹⁰¹.

Durante la medida de desviaciones verticales se deben tener en cuenta una serie de aspectos importantes, por ejemplo, si se quiere medir la magnitud total de una desviación vertical, no se debe dejar al paciente inclinar la cabeza tras el foróptero pues una posición de tortícolis puede disminuir o llegar a eliminar una desviación. Si se cree que la graduación del sujeto puede inducir un efecto prismático vertical significativo se podrían colocar momentáneamente los agujeros estenopéicos para forzar al sujeto a mirar por el centro óptico de las lentes. Esta falta de alineamiento entre el eje visual y el centro óptico de las lentes es una importante fuente de error al medir la desviación vertical que nos puede llevar a medir desviaciones verticales erróneamente elevadas o a ocultar desviaciones verticales reales¹¹⁵. En el estudio presente no fue necesario utilizar los agujeros estenopéicos en ningún caso.

Volviendo a la discusión de aspectos generales relacionados con la repetibilidad, se ha comentado que las medidas de heteroforia **con foróptero** pueden modificar los

hallazgos clínicos respecto a la medida con gafa de pruebas (espacio libre) al haber diferencias en la vergencia proximal, en la posición de cabeza u ojos o en la restricción del campo de visión periférica^{116, 117}. Lam *et al.*¹⁰⁰ compararon las medidas de Thorington modificado y de varilla de Maddox realizadas con foróptero y con gafa de pruebas y vieron que los sujetos tendían a ser menos exofóricos (o más endofóricos) cuando se les corregía con el foróptero, aunque la diferencia no fue significativa. Casillas y Rosenfield¹⁰¹ compararon la repetibilidad de tres pruebas subjetivas de medida de la desviación: von Graefe, varilla de Maddox y Thorington modificado (VL y VP) cuando se aplicaban usando el foróptero *versus* gafa de pruebas. Obtuvieron diferencias significativas en las medidas de la desviación horizontal aunque, a diferencia de lo que ocurría en el estudio de Lam *et al.*¹⁰⁰, en cuatro de las seis situaciones de medida la media tendió hacia exo y no hacia endo al medir con el foróptero. Luego, al medir heteroforias, no ha quedado demostrada la presencia de mayor convergencia proximal al mirar a través del foróptero respecto al espacio libre, propuesta por Lam *et al.*¹⁰⁰. Por otra parte, las técnicas tendieron a mostrar mejor repetibilidad si se aplicaban en el espacio libre en comparación a su aplicación en el foróptero, aunque en ningún caso las diferencias superaron los límites del intervalo de acuerdo al 95% de cada procedimiento¹⁰¹.

Los CDRs reflejados en la Tabla 15 y la Tabla 16 además de permitir comparar las distintas pruebas estudiadas, representan el límite inferior de las variaciones potenciales intra-examinador en la medida de la forias. Es decir, para poder afirmar que se ha producido un cambio real en la medida, la diferencia debe ser al menos igual al CDR correspondiente.

CONCORDANCIA

Los resultados del estudio de concordancia reflejaron que en la medida de la desviación horizontal los CDCs eran notoriamente más elevados si se consideraban las diferencias con signos respecto a las diferencias en valor absoluto (ver Tabla 17). Basándonos en las diferencias con signos la mayoría de las pruebas no podría considerarse como intercambiables pues los CDCs fueron en muchos casos elevados (entre $\pm 2,49\Delta$ y $\pm 7,50\Delta$), especialmente en cerca.

La comparación de estudios que evalúan las distintas pruebas para medir forias es complicada pues en ellos se han usado diferentes procedimientos clínicos o experimentales y es frecuente que hayan realizado análisis estadísticos diferentes, en concreto es muy habitual el uso de coeficientes de correlación en vez de estudios de concordancia (ver Tabla 20).

Cuanto más parecido sea el sistema de disociación utilizado por diferentes métodos de medida, es de esperar que se produzca una mayor concordancia que si se usaran distintos sistemas de disociación. Los resultados encontrados (ver Figura 21) muestran que, considerando los valores medios de forias horizontales, en VL la prueba que tendió a dar mayor endoforia fue Maddox, seguida por von Graefe, punto de neutralización en el cover test y por último la prueba que tendió a reflejar mayor exoforia fue el punto central entre neutralización e inversión en el cover test. En VP la tendencia fue algo diferente, siendo el test de Thorington modificado el que mostró mayor endoforia, seguido del punto neutro al cover test, Maddox, von Graefe y de nuevo como prueba con los resultados más exofóricos el punto central del cover test. Respecto a la tendencia de los valores de las distintas pruebas, los resultados del presente estudio están de acuerdo con los encontrados por otros autores que

comparaban los métodos subjetivos de medida de la desviación aquí estudiados^{96, 101, 103}.

A la vista de las gráficas de concordancia, Figura 25 y Figura 26, se puede ver que la concordancia entre pruebas objetivas y subjetivas disminuye al aumentar la magnitud de la desviación. En estos casos Bland y Altman²⁸ sugieren que se divida el rango de valores de foria en partes y que se estudie la concordancia entre métodos en cada tramo^{††}. Sin embargo, observando las gráficas de la Figura 25 y la Figura 26 se puede ver que la variación del grado de concordancia sucede en un intervalo dióptrico muy estrecho (de unas pocas dioptrías prismáticas) por lo que la separación en distintos tramos de magnitud de desviación no tiene mucho sentido desde el punto de vista clínico. Por ejemplo, al comparar el cover test (punto de neutralización) con von Graefe en VL (gráfica central, izquierda de la Figura 25) habría que separar el rango dióptrico por ejemplo en dos partes que podrían ser: forias $\leq 2\Delta$ y forias $> 2\Delta$. Sería necesario tomar este pequeño intervalo pues que en él ya se produce una fuerte variación del grado de concordancia. Sin embargo, desde el punto de vista clínico y teniendo en cuenta los resultados del estudio de repetibilidad de estas pruebas, no tiene sentido este análisis de la concordancia en tramos con la muestra de sujetos actual. No obstante, sería interesante estudiar este aspecto en futuras investigaciones.

El procedimiento de medida de forias con el test de **Thorington modificado** es similar al seguido con varilla de Maddox, pero en el primero se pide al sujeto que diga el número de la escala graduada que atraviesa la línea roja, mientras que con Maddox el sujeto mira a una luz puntual.

^{††} En el apartado dedicado a la estereoaquidez se puede ver la aplicación de este tipo de estudio de repetibilidad y de concordancia separando el rango de valores de la variable en tramos.

La presencia del estímulo acomodativo en el test de Thorington podría justificar su tendencia a dar valores más endo de foria en cerca. No obstante, Lam *et al.*¹⁰⁰ no encontraron diferencias significativas entre ambos métodos. Estos autores también vieron que en el test de Thorington el uso de prismas para medir la desviación frente al uso de su escala graduada, no afectaba a los resultados de la foria medida.

No es fácil dar una razón por la que la técnica de **von Graefe** tiende a dar valores más exofóricos, pero parece poco probable que se deba a una mala respuesta acomodativa pues se utilizan estímulos relativamente pequeños. Por otra parte, dado que la diferencia entre métodos se presenta tanto en lejos como en cerca, parece poco probable que los resultados se deriven de una vergencia proximal o tónica reducida. Se ha propuesto como posible explicación que al comenzar la medida en la prueba de von Graefe el paciente ve el estímulo a través de un prisma de 12Δ base interna (BI) y los sujetos, incluso en condiciones de disociación, podrían ejercer una respuesta de disparidad de divergencia en un intento de reducir la separación horizontal entre las imágenes en diplopía que llevaría a una adaptación de vergencias¹¹⁸.

En el estudio presente al medir con la técnica de **Maddox**, también se comenzó por un prisma de 12Δ BI, de modo que esperaríamos un comportamiento similar tanto en Maddox como en von Graefe. Esta situación sí se dió en cerca con valores medios de foria prácticamente iguales, $-1,08\Delta$ (Maddox) y $-1,09\Delta$ (von Graefe) respectivamente; sin embargo, en lejos el valor medio de ambas pruebas fue algo diferente, $+3,15\Delta$ (Maddox) y $+0,52\Delta$ (von Graefe). Una posible contribución a este hecho puede provenir de que mientras que en von Graefe en todos los sujetos, incluso en ortofóricos, el punto de partida fue 12Δ BI; en la medida con Maddox en los sujetos que sin prisma vieron el punto y la línea superpuestos (VL: Inicial: 4 sujetos, Final: 3

sujetos; VP: Inicial: 19 sujetos, Final: 12 sujetos) se anotó directamente cero como medida de foria sin proceder a medir partiendo de 12Δ BI.

Respecto al **cover test**, uno de los objetivos de este estudio fue investigar si utilizando diferentes puntos finales de la compensación con prismas en el cover test se llegaba a diferencias significativas en la medida de heteroforias. Aunque pueden aparecer diferencias estadísticamente significativas, como ocurrió en nuestro estudio, sólo se considera clínicamente significativa una diferencia mayor o igual a 2Δ ^{109, 110}. Johns *et al.*⁹¹ también compararon los dos valores del cover test en cerca correspondientes a la primera neutralización y al punto central entre primera neutralización e inversión. Ellos encontraron una diferencia media siempre menor de 2Δ tanto en el promedio con signos como en valor absoluto. En el estudio presente las diferencias promedio con y sin signos en ambas distancias fueron levemente superiores, situándose en todos los casos en torno al mínimo movimiento clínicamente distinguible, 2Δ (desde $1,86\Delta$, con signo VL hasta $2,27\Delta$, valor absoluto VP).

1.15. Conclusiones

A continuación recopilaremos las conclusiones más relevantes de este apartado:

- Todas las técnicas presentaron mejor repetibilidad al medir la desviación en VL, que al medir en VP.
- Tanto en lejos como en cerca, la mejor repetibilidad la presentó el prisma cover test hasta el punto de neutralización.
- Si sólo se consideran las técnicas subjetivas, la prueba que presentó mejor repetibilidad en VP fue la técnica de Thorington modificada.

- La técnica de von Graefe se mostró como la prueba subjetiva menos repetible para medir tanto forias horizontales como verticales.
- Las tres técnicas de medida de forias verticales estudiadas mostraron una buena repetibilidad con DMs y CDRs clínicamente insignificantes.
- La concordancia entre métodos objetivos y subjetivos empeora a medida que aumenta la magnitud de la desviación, no obstante serán necesarios estudios futuros para caracterizar mejor este comportamiento.
- Todos los CDCs aumentaron notoriamente si se consideraban las diferencias con signos respecto a las diferencias en valor absoluto. Si se consideran sólo éstas últimas se puede decir que la concordancia entre pruebas fue bastante buena teniendo en cuenta las importantes diferencias metodológicas existentes.
- Varilla de Maddox tendió a dar resultados más endo (o menos exo) que von Graefe tanto en VL como en VP. En VP el test de Thorington modificado fue la prueba que tendió a reflejar valores más endo de foria.

Reservas fusionales

Introducción

La medida de los rangos de vergencias fusionales es una prueba interesante cuando se está estudiando el estado binocular de un paciente^{58, 119}. El reflejo de fusión es el encargado de mantener compensada una foria, por tanto es interesante conocer qué parte de la amplitud total de vergencias se necesita para compensar una determinada foria. En general, el sistema visual es capaz de mantener un buen rendimiento mientras no se utilicen más de dos tercios de la amplitud total¹²⁰.

El rango total de vergencias se compone de vergencia fusional negativa (VFN) o divergencia, medible con prismas base interna (BI) y de vergencia fusional positiva (VFP) o convergencia, que se puede medir con prismas base externa (BE). Para caracterizar las reservas fusionales se miden tres parámetros: emborronamiento (E), rotura (Ro) y recobro (Re). El emborronamiento es una medida de la cantidad de vergencia fusional relativa (libre de acomodación). La rotura mide la cantidad total de vergencia fusional. Mientras que el recobro proporciona información sobre la capacidad del paciente para recuperar la visión binocular tras haber llegado a la diplopía¹⁴, posiblemente debido a un reflejo de fusión más una componente de vergencia voluntaria¹⁴.

Cuando se miden las reservas fusionales positivas el sujeto normalmente nota que el estímulo de fijación se emborrona antes de que se rompa la fusión y se manifieste la diplopía. Sin embargo, cuando se estimula la divergencia en lejos, la acomodación ya está al mínimo pues el paciente permanece emetropizado durante las medidas y no se

espera que haya punto de emborronamiento[♦]. Al medir las reservas fusionales negativas en cerca sí se suele manifestar la borrosidad de la tarjeta antes de la diplopía pues cuando se fuerza a los ojos a diverger se induce una relajación de la acomodación.

Se pueden encontrar bastantes trabajos sobre la medida de los rangos de vergencias fusionales. Por ejemplo, se ha demostrado que los rangos fusionales están afectados por el orden de aplicación de las medidas de convergencia y divergencia^{121, 122}. En general se admite que cuando se repite la medida de vergencias fusionales en el mismo paciente, el segundo valor es bastante diferente del primero⁷. Sin embargo, son escasas las investigaciones que hayan estudiado la repetibilidad de las diferentes técnicas y la equivalencia entre ellas^{7, 123-125}. Además, las investigaciones encontradas presentan distintas limitaciones. Por ejemplo, sus resultados se basaron en medidas obtenidas en una única sesión o los análisis estadísticos utilizados no eran los recomendados, o si lo eran, las medidas se centraban sólo en el estudio de la rotura BE⁷. Varias de las investigaciones también tenían otros objetivos, y el estudio de la repetibilidad sólo se abordaba indirectamente o de forma secundaria^{123, 124, 126}.

Diasporámetro y barra de prismas son los métodos habituales para evaluar la amplitud de la reserva fusional. Mediante los diasporámetros, la amplitud de vergencias es medida a través de un incremento suave y progresivo de potencia prismática en comparación con la barra de prismas donde los incrementos de potencia son discretos y a saltos. Se ha visto que la medida de la amplitud de vergencias con diasporámetros proporcionan resultados bastante repetibles en adultos jóvenes¹²⁴. También se ha visto que los resultados son menos repetibles en niños y es una prueba con una variabilidad

[♦] Algunos sujetos pueden apreciar cierto grado de borrosidad cuando se estimula la divergencia en lejos, situación normalmente asociada a casos donde la acomodación no está totalmente relajada y hace sospechar de la presencia de una hipermetropía latente o de una hipercorrección miópica.

inter-examinador alta⁷. La medida de los rangos de vergencia en el espacio libre mediante barra de prismas se parece más a las condiciones habituales y es especialmente útil al medir a niños pequeños pues permite ver los ojos del paciente y controlar de forma objetiva la fiabilidad de las respuestas del sujeto.

El punto próximo de convergencia (PPC) es el punto donde se cruzan los ejes visuales cuando se está llevando a cabo el máximo esfuerzo de convergencia y aun se consigue mantener visión binocular única. La medida del PPC es un aspecto importante en un examen visual rutinario porque constituye una de las principales herramientas diagnósticas de la insuficiencia de convergencia y ayuda a controlar las mejorías logradas tras su tratamiento¹⁴. Los valores normales para el punto de rotura de la convergencia tradicionalmente se han considerado entre 8 y 10 cm según los manuales optométricos; aunque varios estudios lo han situado pocos centímetros por delante de la cara del paciente¹²⁷.

La medida del PPC es una prueba rápida y fácil de llevar a cabo, no requiere equipamiento especial y según Rouse *et al.* aporta resultados muy repetibles⁷. Se han propuesto distintos métodos de medida del PPC¹²⁸ y Ciuffreda¹²⁹ ha señalado que los tests de fijación con alta demanda acomodativa resultan en un PPC más cercano, posiblemente debido a la mayor convergencia acomodativa estimulada. No obstante, las diferencias respecto a otros estímulos de fijación como la punta del dedo o la de un lapicero no fueron clínicamente significativas¹³⁰. La convergencia casi siempre se inicia reflexivamente pero como Adler¹³¹ sugirió, el PPC también está fuertemente influenciado por el nivel de atención del sujeto. Otros aspectos que también influyen son el esfuerzo del paciente para converger (convergencia voluntaria), el nivel de acomodación estimulado (convergencia acomodativa), la agudeza visual de cada ojo y posiblemente también el tipo de estímulo de fijación¹³. Algunas de estas variables son

difíciles de controlar, lo que podría influir negativamente en la repetibilidad de la prueba.

Uno de los objetivos de esta parte del estudio fue establecer el grado de repetibilidad de la medida de los rangos de vergencia fusional, realizada con diasporámetro y con barra de prismas, así como analizar la concordancia entre ambas técnicas. Por otra parte, también se quiso estudiar la repetibilidad de la medida del PPC y analizar la correlación con la reserva fusional positiva medida con diasporámetro y con barra de prismas.

Método

Se tomaron medidas en dos ocasiones diferentes del punto próximo de convergencia (PPC) (rotura y recobro) y también de los rangos de vergencia fusional negativa (VFN) y positiva (VFP) (emborronamiento, rotura y recobro) en lejos y en cerca, con diasporámetro y con barra de prismas. Para las medidas de los rangos de vergencia fusional, en primer lugar se midió el rango de vergencia en lejos y después en cerca y, para cada una de las dos distancias, siempre se midió primero la VFN y después la VFP. Esto se debió a que la adaptación prismática en BE parece ser mayor que la que se puede dar en BI, de modo que se podría producir una reducción de la amplitud de vergencia medida en BI si se midiera después, y no antes, de la de BE¹³. Como la medida de los rangos de vergencia puede ser afectada por el fenómeno de adaptación prismática¹³²⁻¹³⁴, se dio un periodo de descanso de unos 15 segundos entre cada medida. Durante este periodo los sujetos miraban en visión lejana (6 m) para minimizar los efectos de dicha adaptación prismática⁹⁴. El orden de las pruebas (diasporámetro, barra de prismas y PPC) fue establecido aleatoriamente para evitar que los efectos del aprendizaje y/o la fatiga del sujeto afectaran a los resultados.

El estímulo de fijación fue común y consistió en una columna de letras correspondiente a una AV decimal de Snellen de 0,8. Para las medidas de lejos el estímulo se proyectó a 6 m y para las de cerca se presentó en una tarjeta situada a 40 cm. Para las medidas de lejos se usó la iluminación de sala y se añadió un foco de luz dirigido hacia la tarjeta para las medidas en cerca, asegurándonos que no se produjeran sombras sobre la tarjeta de fijación.

Durante las medidas el sujeto debía esforzarse por mantener las letras del test de fijación tan nítidas como le fuera posible y se le pidió que nos avisara en el momento de apreciar la primera borrosidad mantenida del test (emborronamiento), en el momento en que se desdoblara la columna de letras sin poder volver a fusionarlas (rotura) y cuando de nuevo volvía a ver una única columna de letras (recobro). Para facilitar la detección del punto de emborronamiento, antes de comenzar con las medidas se simuló la sensación de borrosidad con un +0,50 D sobre la compensación de lejos. Además, puesto que el criterio para determinar el emborronamiento fue el de la primera borrosidad mantenida, cuando el sujeto manifestó ver borroso se cesó momentáneamente el aumento de la potencia prismática y se le pidió que nos confirmara si efectivamente no podía aclarar el test. Se consideró como punto de rotura la menor potencia prismática, o la mayor distancia de fijación en el PPC, que hacía que el sujeto viera doble de forma mantenida, es decir, se buscaba la amplitud máxima de vergencias. Una vez que el estímulo se veía doble, se aumentaba la demanda prismática unas 4-5 Δ (o 1-2 cm en el PPC) y, posteriormente, se reducía de nuevo la demanda hasta que el estímulo se volviera a ver simple de nuevo (punto de recobro).

A continuación se darán algunos detalles más sobre la metodología seguida en cada una de las pruebas:

VERGENCIAS FUSIONALES CON LOS DIASPORÁMETROS DEL FORÓPTERO

Los sujetos llevaron su compensación de lejos colocada en el foróptero y se ajustó adecuadamente la distancia interpupilar. La demanda prismática se introdujo mediante los prismas de Risley colocados ante ambos ojos (vergencias suaves). Partiendo de cero se fue aumentando simétricamente la potencia prismática en cada ojo (ver Figura 28).



Figura 28. Medida de las vergencias fusionales horizontales con diasporámetro

VERGENCIAS FUSIONALES CON BARRA DE PRISMAS

Los sujetos llevaron su compensación de lejos colocada en gafa de pruebas. Durante las medidas, se situó una barra de prismas horizontal estándar ante el ojo derecho del sujeto (los pasos de prisma eran 1, 2, 4-20 en pasos de 2 Δ y 25-40 en pasos de 5 Δ) y se aumentó progresivamente la potencia prismática a un ritmo de un paso cada 2 o 3 segundos (ver Figura 29) (vergencias a saltos). Como es relativamente fácil inducir un efecto prismático vertical indeseable debido a una inclinación no intencionada de la barra de prismas, fue esencial controlar cuidadosamente la orientación de la barra de prismas, especialmente en altas potencias. Aunque se espera que los sujetos sean capaces de apreciar el punto de borrosidad al medir la VFP en VL y VP y al medir VFN

en VP, en la práctica a muchos sujetos les cuesta apreciar la borrosidad del test debido a que los pasos de vergencia son de varias dioptrías prismáticas.



Figura 29. Medida de las reservas fusionales horizontales con barra de prismas

PUNTO PRÓXIMO DE CONVERGENCIA

Durante la medida del PPC el sujeto permaneció sentado y con su graduación de lejos en gafa de pruebas utilizando lentes de aro grande. El examinador se situó frente al sujeto para poder ver simultáneamente sus dos ojos y le instruyó adecuadamente.

La tarjeta se colocó inicialmente a 50 cm sobre una regla de acomodación¹³⁵ (ver Figura 30) y se acercó lentamente hacia la raíz nasal del sujeto a unos 2 cm/s^δ. Esta técnica requería que el sujeto indicara al examinador el punto en el que el test de fijación se veía doble. En ese momento el sujeto debía esforzarse por volver a recuperar la visión simple. Si lo lograba se continuaba acercando la tarjeta hacia los ojos del paciente hasta la primera diplopía mantenida (rotura), es decir, cuando el

^δ Un movimiento demasiado rápido de la tarjeta tiende a hiperestimar el PPC, mientras que si se mueve demasiado despacio (o comenzando desde muy lejos) el sujeto podría aburrirse y dar un PPC artificialmente reducido de la medida

sujeto no lograba fusionar de nuevo. Posteriormente se alejaba progresivamente la tarjeta y se pidió al sujeto que se esforzara por recuperar la visión simple lo antes posible (recobro). Las distancias de rotura y recobro se midieron respecto al plano de la gafa con una precisión del 0,5 cm. En aquellos sujetos que lograron mantener la fusión hasta que la tarjeta alcanzó la raíz nasal, se anotó cero como valor de la rotura y no se midió el recobro.

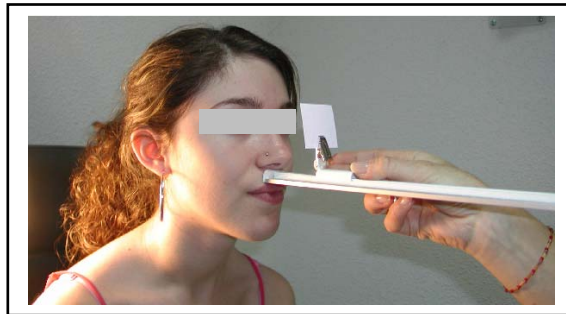


Figura 30. Medida del punto próximo de convergencia (PPC).

1.16. Resultados

REPETIBILIDAD

Respecto al PPC, la Figura 31 muestra las gráficas de Bland y Altman correspondientes a las medidas inicial y final de rotura (N=61) y de recobro (N=39, pues no se puede obtener esta medida en aquellos sujetos que consiguen mantener la fusión hasta la raíz nasal). Se puede ver que las DMs tanto para la rotura como para el recobro fueron estadística y clínicamente insignificantes (ver Tabla 21). Los CDRs también fueron clínicamente buenos (cerca de ± 3 cm) teniendo en cuenta las características de la prueba. En resumen, se puede decir que la medida del PPC es una prueba muy repetible.

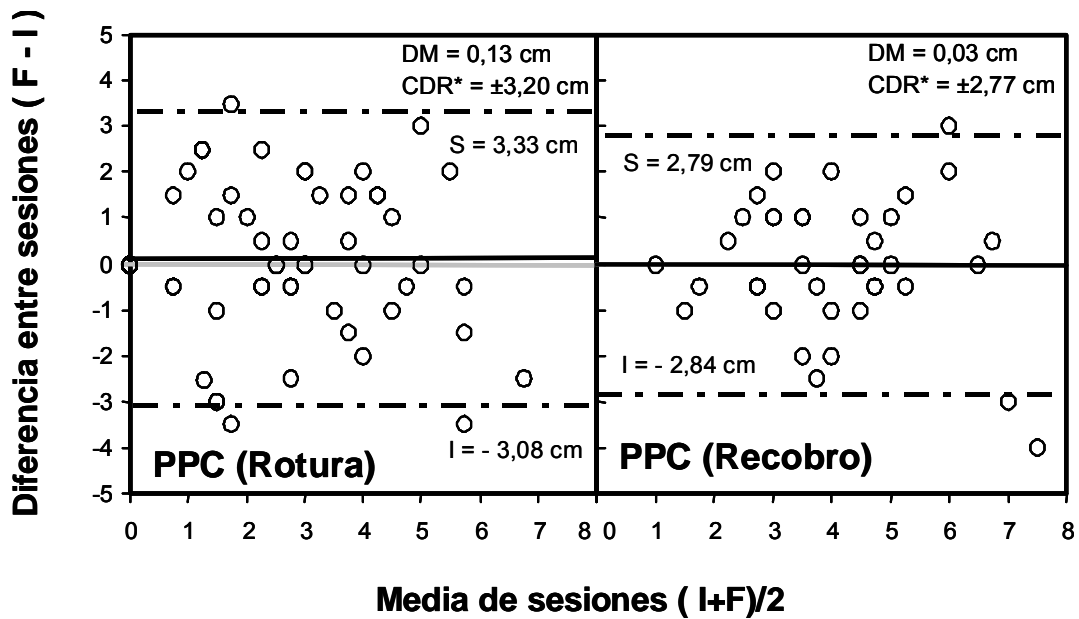


Figura 31. Diferencia media (DM), límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) y coeficiente de repetibilidad (CDR) para la medida del punto próximo de convergencia (PPC). Tanto en la rotura (Ro) como en el recobro (Re) la distribución de las diferencias no fue normal y se tomó como coeficiente de repetibilidad el percentil 95 del valor absoluto de las diferencias (CDR*).

Tanto para diaporámetro, como para barra de prismas las diferencias medias (DMs) y los intervalos entre los límites de acuerdo al 95% (CDRs) han sido menores, es decir, la repetibilidad ha sido mayor, en las medidas de reservas fusionales BI (divergencia) que en las de BE (convergencia), tanto en lejos como en cerca.

Tabla 21. Repetibilidad en la medida del PPC

	Media (cm)	DE (cm)	DM (F-I) (cm)	p (test-t)	CDR (cm)
Rotura	2,51	1,98	0,13	p = 0,6	±3,20*
Recobro	4,08	1,87	0,03	p = 0,9	±2,77*

Clave: DE = desviación estándar. DM = diferencia media. CDR = coeficiente de repetibilidad. F = final, I = inicial. cm = centímetros.

* El CDR aparece con un asterisco indicando que se ha sustituido éste por el percentil 95º del valor absoluto de las diferencias

La Tabla 22 resume los resultados del estudio de la repetibilidad de la medida de los rangos de vergencias horizontales con diaporámetro y con barra de prismas. Las DMs entre sesiones de medida encontradas para los puntos de rotura y recobro de VFN fueron tan pequeñas que pueden considerarse clínicamente insignificantes (menores de $0,5\Delta$ en todos los casos), mientras que las diferencias encontradas en los rangos de VFP entre la sesión inicial y la final normalmente superaron las 2Δ . Los CDRs también fueron algo mayores para la medida de VFP que en la de VFN.

Los resultados para las medidas de las reservas fusionales con barra de prismas se muestran en la Figura 32 (6 m) y en la Figura 35 (40 cm), y para las medidas con diaporámetro en la Figura 33 (6 m) y en la Figura 35 (40 cm). En estas gráficas, el eje-y corresponde a la diferencia con signos entre las medidas de ambas sesiones (Final – Inicial) y el eje-x corresponde a la media de las dos medidas (en Δ). Si la prueba muestra buena repetibilidad, la DM será cercana a 0 y el intervalo correspondiente a $DM \pm 1,96 \cdot DE$ de las diferencias, o límites de acuerdo al 95% será estrecho. Se puede ver que ninguna de las gráficas muestra tendencia a que la diferencia aumente con el valor prismático, es decir, la repetibilidad de los tests no cambia con el rango fusional.

Tabla 22. Repetibilidad de la medida de las reservas fusionales horizontales con barra de prismas y con los diasporámetros del foróptero.

MEDIDA		BARRA DE PRISMAS				
		Media (Δ)	DE (Δ)	DM F-I (Δ)	p(test t)	CDR (Δ)
VISIÓN LEJANA	Ro BI	8,63	1,94	0,20	0,5	$\pm 4,00^*$
	Re BI	6,26	1,82	0,36	0,1	$\pm 4,00^*$
	E BE	12,92	5,22	1,17	0,3	$\pm 9,73$
	Ro BE	23,25	7,68	3,85	<0,0001	$\pm 13,93$
	Re BE	14,50	4,17	2,44	<0,0001	$\pm 8,30^*$
VISIÓN PRÓXIMA	E BI	8,75	3,37	1,54	0,1	$\pm 6,60^*$
	Ro BI	12,14	3,35	-0,08	0,9	$\pm 8,05^*$
	Re BI	9,78	3,02	-0,21	0,6	$\pm 6,00^*$
	E BE	17,08	6,45	2,16	0,05	$\pm 10,21$
	Ro BE	28,91	9,09	1,30	0,2	$\pm 15,00^*$
	Re BE	19,65	5,98	2,22	0,004	$\pm 11,50^*$
MEDIDA		DIASPORÁMETRO				
		Media (Δ)	DE (Δ)	DM F-I (Δ)	p(test t)	CDR (Δ)
VISIÓN LEJANA	Ro BI	9,99	2,36	0,05	0,9	$\pm 4,82$
	Re BI	5,33	2,02	0,47	0,03	$\pm 3,00^*$
	E BE	15,69	6,36	-0,54	0,7	$\pm 14,00^*$
	Ro BE	24,68	7,35	2,97	0,0002	$\pm 11,29$
	Re BE	11,80	5,68	2,63	0,0003	$\pm 12,00^*$
VISIÓN PRÓXIMA	E BI	11,73	3,80	0,39	0,7	$\pm 6,43$
	Ro BI	15,98	4,29	-0,05	0,9	$\pm 7,00^*$
	Re BI	8,24	4,04	-0,44	0,4	$\pm 7,49$
	E BE	22,08	6,83	1,17	0,4	$\pm 8,83$
	Ro BE	29,24	8,36	2,08	0,0001	$\pm 7,74$
	Re BE	19,22	6,78	4,02	<0,0001	$\pm 16,40^*$

Clave: E = emborronamiento, Ro = rotura, Re = recobro, VFN/VFP = vergencia fusional negativa/positiva. DE = desviación estándar. DM = diferencia media, CDR = coeficiente de repetibilidad, Δ = dioptría prismática, d = diasporámetro, b = barra de prismas. F = final, I = inicial

* En los casos donde CDR aparece con un asterisco, se ha sustituido éste por el valor del percentil 95° del valor absoluto de la diferencia

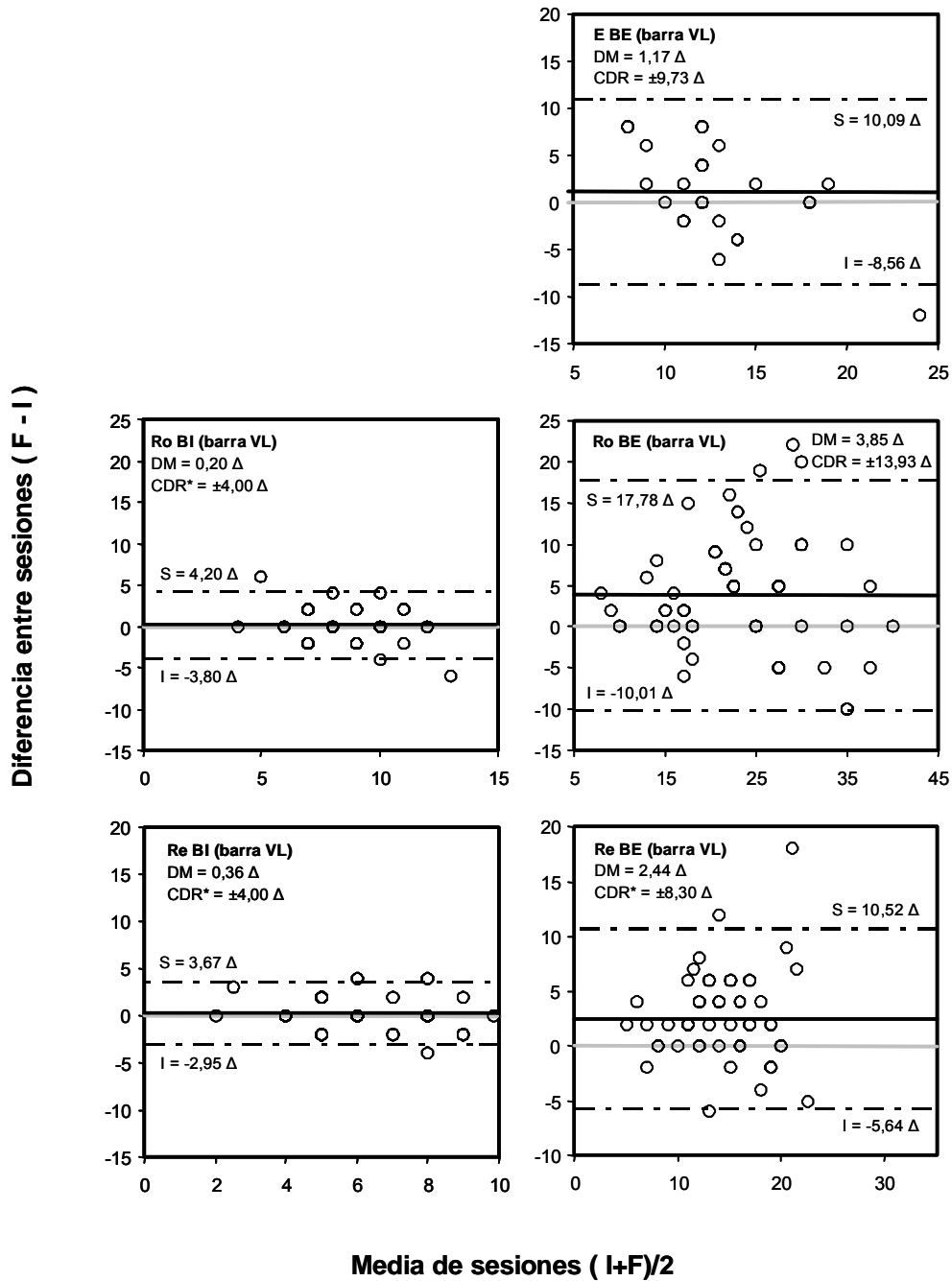


Figura 32. Diferencia media (DM), límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) y coeficiente de repetibilidad (CDR) para la medida de las reservas fusionales (VFN y VFP) con barra de prismas en lejos. Cuando la distribución de las diferencias no fue normal y se tomó como coeficiente de repetibilidad el percentil 95 del valor absoluto de las diferencias (CDR*).

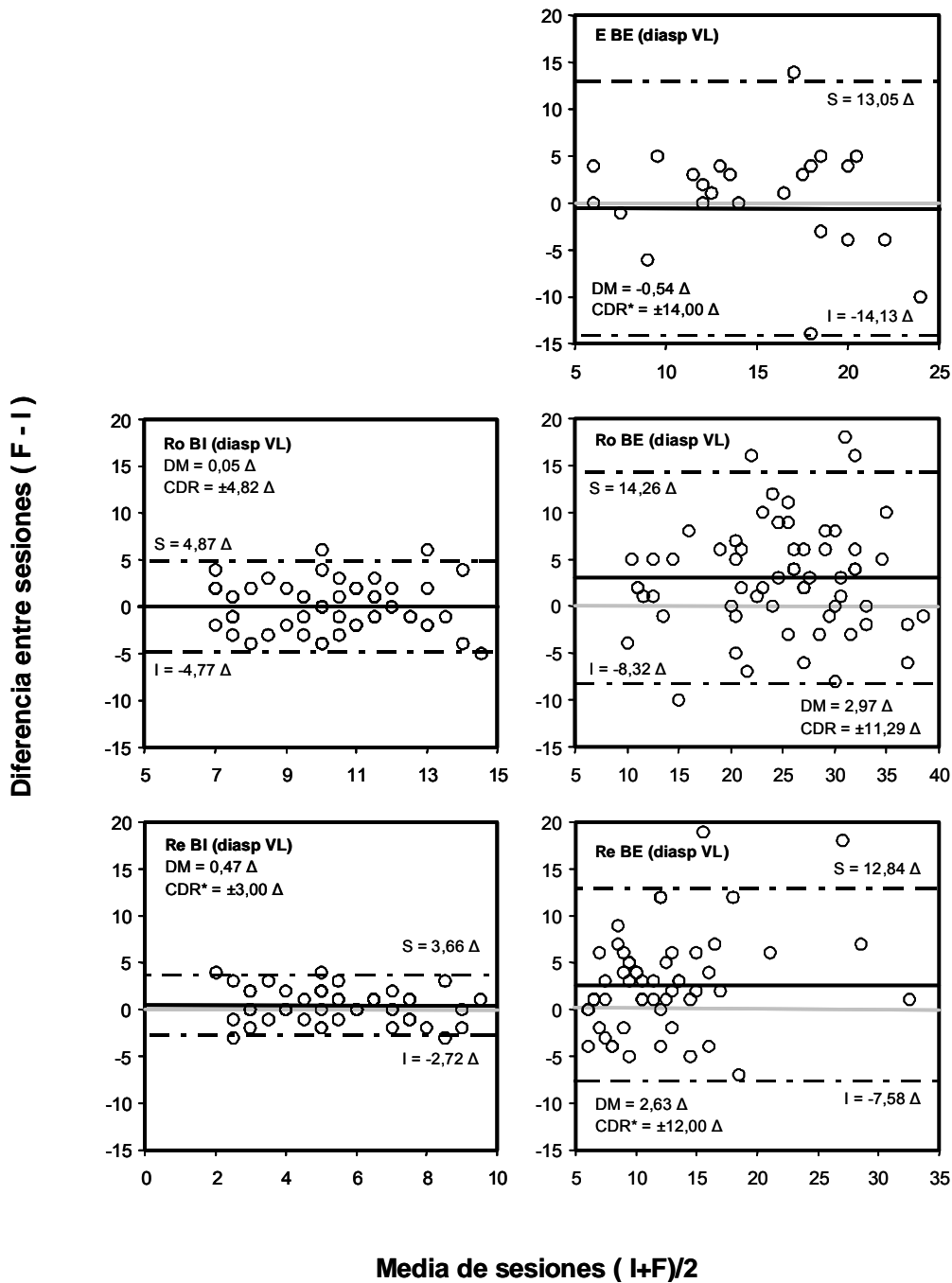


Figura 33. Diferencia media (DM), límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) y coeficiente de repetibilidad (CDR) para la medida de las reservas fusionales (VFN y VFP) con diaspórametro en lejos. Cuando la distribución de las diferencias no fue normal y se tomó como coeficiente de repetibilidad el percentil 95 del valor absoluto de las diferencias (CDR*).

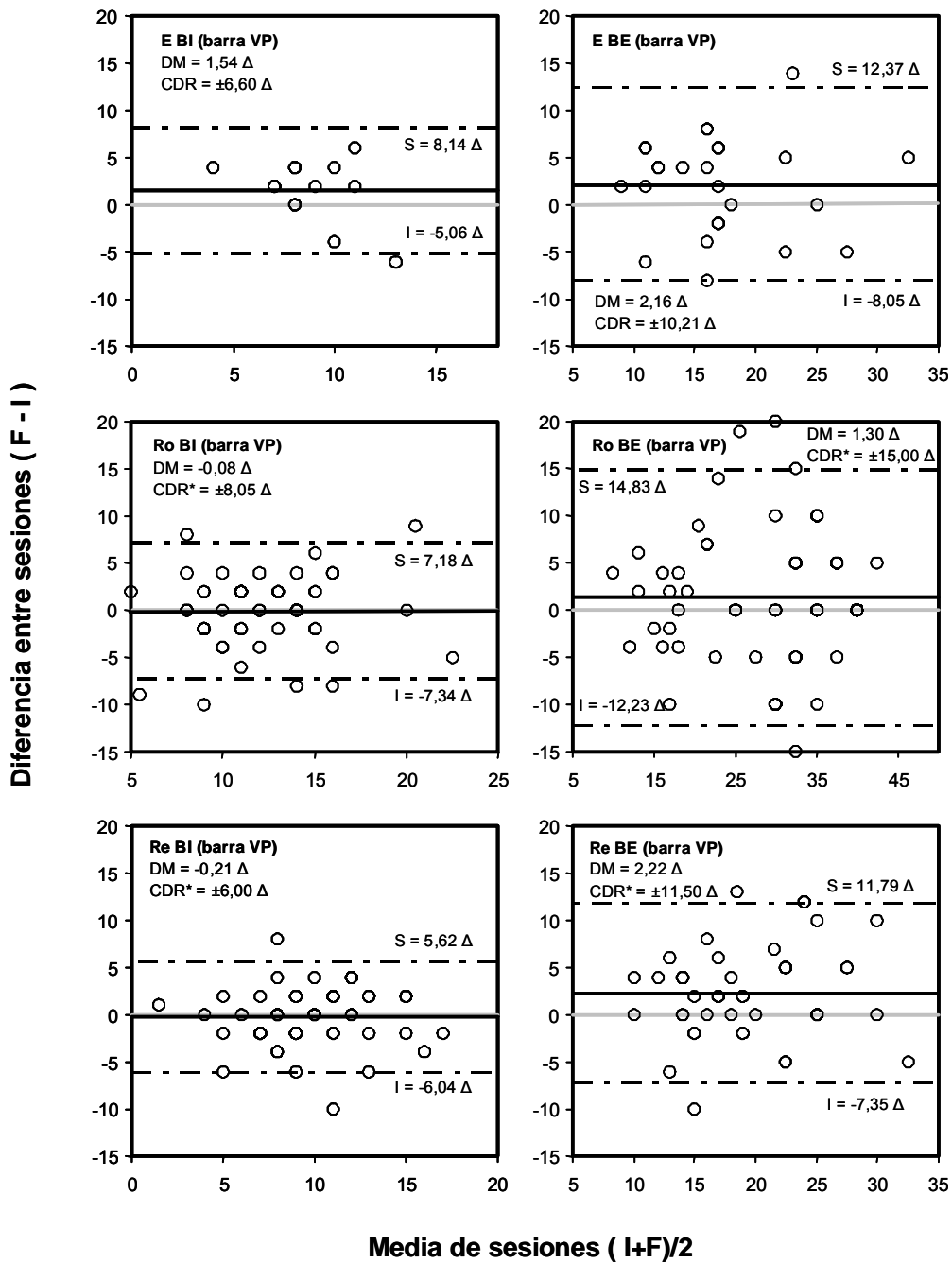


Figura 34. Diferencia media (DM), límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) y coeficiente de repetibilidad (CDR) para la medida de las reservas fusionales (VFN y VFP) con barra de prismas en cerca. Cuando la distribución de las diferencias no fue normal y se tomó como coeficiente de repetibilidad el percentil 95 del valor absoluto de las diferencias (CDR*).

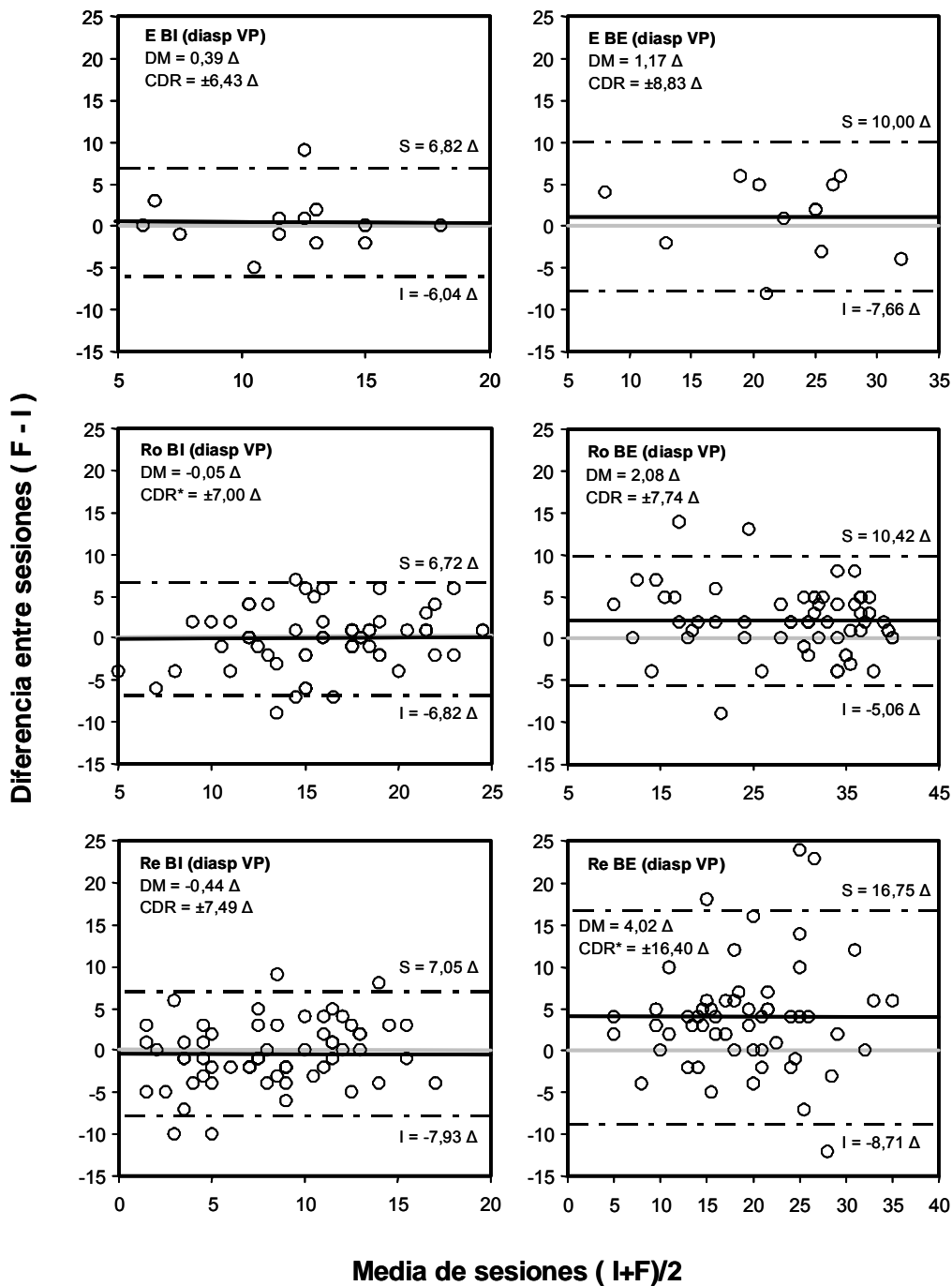


Figura 35. Diferencia media (DM), límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) y coeficiente de repetibilidad (CDR) para la medida de las reservas fusionales (VFN y VFP) con diaspómetro en cerca. Cuando la distribución de las diferencias no fue normal y se tomó como coeficiente de repetibilidad el percentil 95 del valor absoluto de las diferencias (CDR*).

CONCORDANCIA

La Tabla 23 muestra los resultados obtenidos para el estudio de la concordancia entre métodos, las DMs entre las medidas con diaporámetro y con barra de prismas reflejan que los puntos de emborronamiento y rotura se mostraron más altos para las medidas con diaporámetro, mientras que los recobros fueron mayores para la medida con barra de prismas, encontrando diferencias estadísticamente significativas en prácticamente todas las medidas realizadas, tanto en lejos como en cerca. Es decir, se apreciaba antes la borrosidad del estímulo y se rompía antes la fusión con barra de prismas que con diaporámetro, pero se recuperaba antes la fusión con barra de prismas.

En términos clínicos, la concordancia entre ambas técnicas fue moderada, pues aunque las DMs no superaron las $5,05\Delta$, los CDCs fueron amplios si se comparan con los valores esperados para la VFP y VFN con estas técnicas^{136, 137}. La menor DM se encontró para la rotura BE (VFP) en cerca ($DM=0,33\Delta$, $p=0,7$). No obstante es precisamente para este caso donde se obtuvo uno de los intervalos de acuerdo al 95% más abultado ($CDC=\pm 11,18\Delta$). Los emborronamientos también mostraron unas DMs y unos CDCs elevados, de hecho la mayor DM se obtuvo para los valores de emborronamiento BE en cerca donde las medidas con diaporámetro tendieron a ser notoriamente mayores que las medidas con barra de prismas ($DM=5,05\Delta$, $p=0,0004$). No obstante, hay que tener en cuenta que no todos los sujetos apreciaron la borrosidad de la imagen antes de llegar al punto de rotura ($N=40$ en VFP VL, $N=22$ en VFN VP y $N=29$ en VFP VP), lo que puede influir sobre la potencia de los resultados para emborronamiento.

En las gráficas Bland y Altman para el estudio de concordancia entre métodos, el eje-x corresponde a la media entre métodos y el eje-y a su diferencia. Las gráficas para

nuestros resultados se muestran en la Figura 36 (para 6 m) y en la Figura 37 (para 40 cm). Las líneas horizontales representadas en estas gráficas indican la diferencia media (DM) y los límites de acuerdo entre ambos métodos, $DM \pm 1,96 \cdot DE$ de las diferencias (CDC). En estas gráficas se puede ver que los CDCs, especialmente para las medidas de VFP, son suficientemente elevados como para poder decir que los dos métodos no son intercambiables.

Tabla 23. Concordancia de la medida de vergencias fusionales con diaporámetro y con barra de prismas

MEDIDA	VISIÓN LEJANA			VISIÓN PRÓXIMA		
	DM d-b(Δ)	p (t-test)	CDC (Δ)	DM d-b(Δ)	p (t-test)	CDC (Δ)
E VFN	---	---	---	3,66	0,0001	$\pm 8,00^*$
Ro VFN	1,24	<0,0001	$\pm 3,54$	3,84	<0,0001	$\pm 6,81$
Re VFN	-0,99	0,0001	$\pm 3,71$	-1,54	0,001	$\pm 7,02$
E VFP	2,99	0,0004	$\pm 9,48$	5,05	0,0004	$\pm 13,19$
Ro VFP	1,39	0,04	$\pm 10,06$	0,33	0,7	$\pm 11,18$
Re VFP	-2,61	<0,0001	$\pm 10,50^*$	-1,88	0,006	$\pm 9,00^*$

Clave: E = emborronamiento, Ro = rotura, Re = recobro, VFN/VFP = vergencia fusional negativa/positiva. DM = Diferencia media, CDC = coeficiente de concordancia, Δ = dioptría prismática, d = diaporámetro, b = barra de prismas.

* En los casos donde CDC aparece con un asterisco, se ha sustituido éste por el valor del percentil 95º del valor absoluto de la diferencia

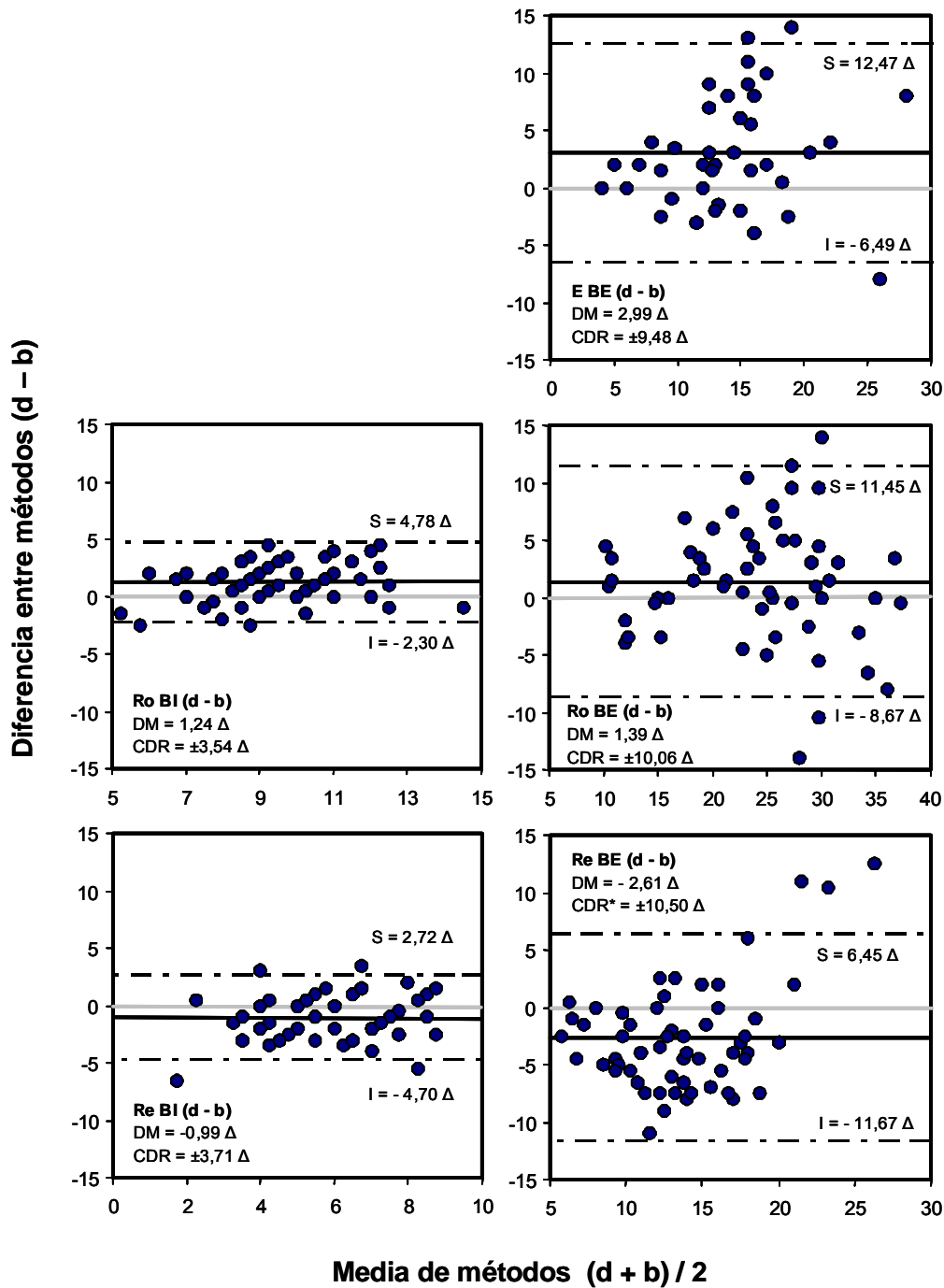


Figura 36. Gráficas de Bland y Altman para el estudio de la concordancia entre métodos que miden las reservas fusionales horizontales en lejos (6 m). La línea continua representa la diferencia media (DM) entre las medidas obtenidas con diaporámetro y con barra de prismas. Las líneas discontinuas indican los límites superior (S) e inferior (I) del intervalo de acuerdo al 95% ($DM \pm 1,96 \cdot DE$). Cuando la distribución de las diferencias no fue normal, se tomó como coeficiente de concordancia el percentil 95 del valor absoluto de las diferencias (CDC*).

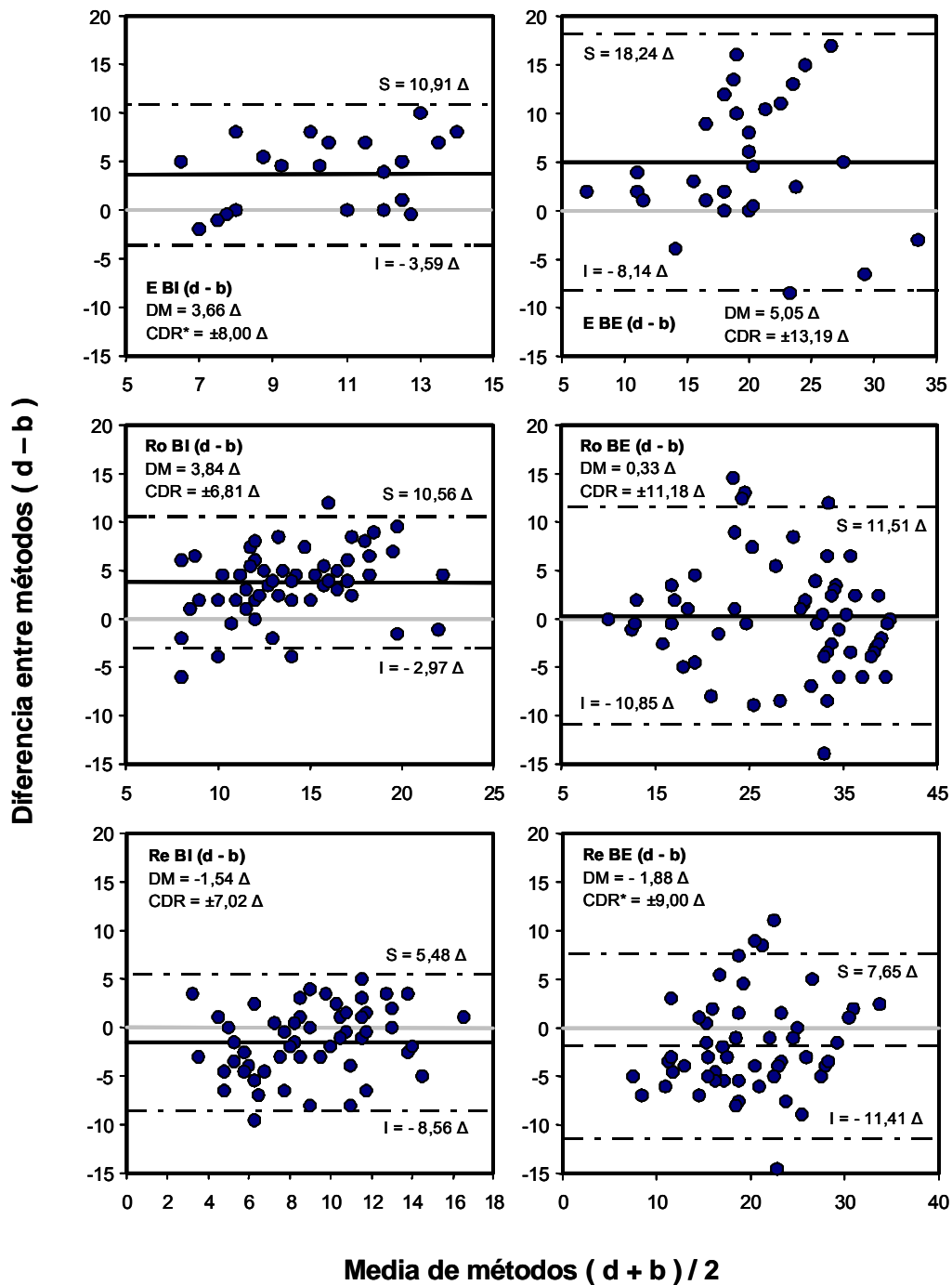


Figura 37. Gráficas de Bland y Altman para el estudio de la concordancia entre métodos que miden las reservas fusionales horizontales en cerca (40 cm). La línea continua representa la diferencia media (DM) entre las medidas obtenidas con diasporámetro y con barra de prismas. Las líneas discontinuas indican los límites superior (S) e inferior (I) del intervalo de acuerdo al 95% ($DM \pm 1,96 \cdot DE$). Cuando la distribución de las diferencias no fue normal, se tomó como coeficiente de concordancia el percentil 95 del valor absoluto de las diferencias (CDC*).

No podemos hacer un análisis de concordancia de Bland y Altman entre la convergencia máxima (punto de rotura) medida con prismas y el PPC puesto que se miden en distintas unidades. Sin embargo, sí se ha realizado un análisis de correlación entre el PPC y la VFP en cerca obtenida tanto con barra de prismas como con diaporámometro. La Figura 38 muestra una baja correlación entre las medidas de rotura en cerca obtenida con la prueba del PPC y mediante la introducción de demanda de convergencia con barra de prismas o con diaporámometro. En la parte inferior de cada gráfica se muestra la ecuación correspondiente al ajuste lineal por mínimos cuadrados y el valor del coeficiente de correlación de Pearson (r) que corroboran la baja correlación entre pruebas.

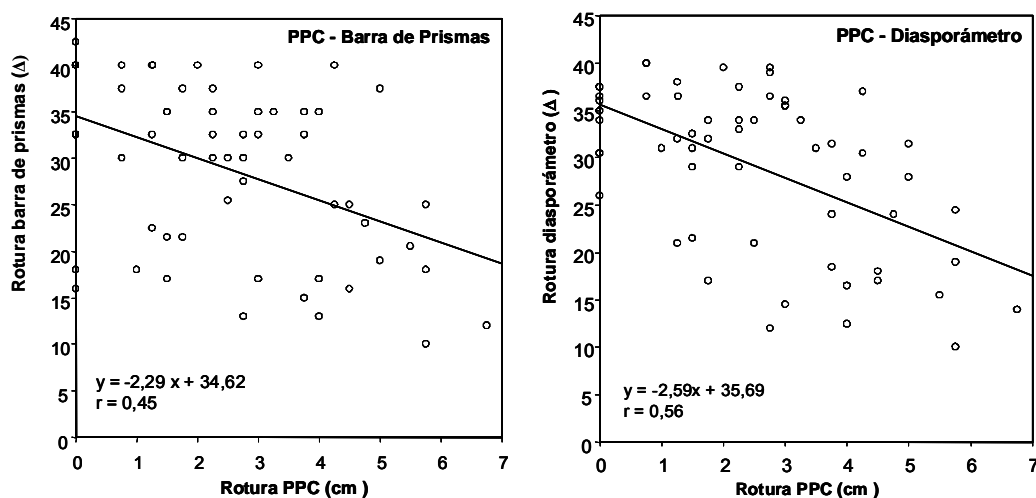


Figura 38. Correlación entre la medida del PPC y la VFP en VP medida con barra de prismas (izquierda) y con diaporámometro (derecha).

1.17. Discusión

REPETIBILIDAD

Sheedy y Saladin¹³⁸ sugirieron que no es raro que se de una diferencia de 10Δ entre una medida de la amplitud de vergencia y otra a menos que se apliquen controles

rigurosos al aplicar las pruebas. Es difícil comparar los resultados del presente trabajo con los de otros estudios que han examinado la repetibilidad de la medida de vergencias a través de varias sesiones de medida^{7, 123-125} debido a que, excepto el trabajo de Rouse *et al.*⁷, no se ha encontrado ningún estudio de repetibilidad que utilice los métodos estadísticos actualmente recomendados. Ello obliga a centrar la discusión de resultados en aspectos fundamentalmente cualitativos.

Rouse *et al.*⁷ evaluaron la rotura Be con diaporómetro en dos sesiones separadas una semana entre sí a 20 niños de entre 10 y 11 años y encontraron que se necesitaría una diferencia intra-examinador (CDR) de 12Δ a 14Δ para indicar una mejoría asociada a determinado tratamiento.

Los resultados de repetibilidad de la medida de vergencias reflejados en la Tabla 22 muestran que, para las medidas con diaporómetro, la variabilidad fue menor en las medidas BI (VFN) que en las medidas BE (VFP). Así, nuestros menores intervalos correspondieron al recobro y la rotura BI en lejos ($\pm 3,00\Delta$ y $\pm 4,82\Delta$ respectivamente), mientras que el mayor intervalo ($\pm 16,40\Delta$) se encontró para el recobro BE en cerca, seguido de los intervalos obtenidos para el emborronamiento, recobro y rotura BE en lejos ($\pm 14,00\Delta$, $\pm 11,29\Delta$ y $\pm 12,00\Delta$ respectivamente). Un comportamiento cualitativo similar fue encontrado por Penisten *et al.*¹²⁴ que estudiaron una muestra de 8 sujetos de 28,9 años de edad media (DE=7,0 años). Midieron los rangos de vergencias con diaporómetro en 4 sesiones diferentes, obteniendo unas desviaciones estándar promedio entre $1,1\Delta$ y $2,3\Delta$ para la BI y entre $1,7\Delta$ y $2,8\Delta$ en la BE. La variabilidad fue siempre mayor en la medida BE que en su correspondiente de BI con la única excepción de la rotura en visión próxima.

Más recientemente, Ciuffreda *et al.*¹²⁵ evaluaron semanalmente durante 10 semanas consecutivas los rangos de vergencias en cerca tanto en el espacio libre (barra de

prismas) como en el foróptero (diasporámetro). Midieron a tres sujetos adultos (los autores) de 26, 31 y 57 años respectivamente. Estudiaron tanto la variabilidad intra-sesión, como la variabilidad inter-sesión. Hubo importantes diferencias metodológicas respecto al presente estudio, por ejemplo, su muestra estuvo formada sólo por 3 sujetos experimentados y en cada sesión tomaban tres medidas de cada parámetro y calculaban su valor promedio para estudiar la repetibilidad inter-sesión. Aunque los propios autores plantearon que probablemente sus resultados representasen la mejor situación de medida, aun así la variabilidad de sus resultados fue parecida a la obtenida por Penisten et al.¹²⁴ y no muy alejada de la de este estudio. Ellos obtuvieron unas desviaciones estándar promedio entre 1,5 Δ y 5,7 Δ para barra de prismas en VP y entre 2,3 Δ y 3,9 Δ para diasporámetro en VP y en el presente estudio las desviaciones estándar se situaron entre 1,8 Δ y 9,1 Δ (ver Tabla 22).

El hecho de encontrar menores intervalos de acuerdo en las medidas BI que en BE puede ser debido a que, como ha sido postulado previamente, las funciones de vergencia BI y BE tengan mecanismos neurales subyacentes diferentes, que justificaría, por ejemplo, que una lesión cerebral pueda producir parálisis de una función pero no de la otra^{125, 139}.

Comparando los resultados de VL con los de VP se observa que tanto para barra de prismas como para diasporámetro en todos los casos, excepto en emborronamiento y rotura BE con diasporámetro, la variabilidad fue menor al medir en lejos. En el estudio de Penisten et al.¹²⁴ también se produjo este fenómeno con la única excepción de la rotura BE. El hecho de que dos de los factores que pueden afectar a la repetibilidad de las medidas de los rangos de vergencias, la acomodación y la convergencia proximal, se mantengan al mínimo durante las medidas en visión lejana podría justificar la menor variabilidad encontrada cuando las medidas se realizaron en lejos.

Punto próximo de convergencia

Rouse *et al.*⁷ examinaron la repetibilidad del PPC (sólo rotura) en niños y encontraron un CDR algo más ancho (examinador 1: $\pm 5,33$ cm, examinador 2: $\pm 5,00$ cm) al hallado en el estudio presente ($\pm 3,46$ cm).

En la medida del PPC, los valores medios obtenidos para el punto de rotura (Media = 2,51 cm; DE = 1,98) y el de recobro (media= 4,08 cm; DE = 1,87) fueron muy similares a los encontrados por Scheiman *et al.*¹⁴⁰ para estímulo acomodativo y una muestra de similares características a las del presente estudio (Rotura: Media = 2,49 cm, DE = 1,74; Recobro: Media = 4,35 cm; DE = 2,74). Tal y como se esperaba el punto de recobro se situó más lejos que el de rotura, con una diferencia entre ellos similar a los encontrados en otros trabajos^{130, 140}. La desviación estándar encontrada en el presente trabajo fue similar a la resultante de trabajos previos^{130, 140} en los que los participantes fueron adultos conocedores de la prueba, de los que cabría esperar unos resultados más consistentes. Esto lleva a pensar que la posible anticipación de vergencias que a menudo se da cuando el estímulo se mueve de forma predecible¹⁴¹ no influye de forma significativa en la variabilidad de los resultados obtenidos.

CONCORDANCIA

Las pruebas de vergencias suaves (diasporámetro) y vergencias a saltos (barra de prismas) están diseñadas para evaluar la amplitud de vergencias fusionales. Varios estudios^{137, 142} han demostrado que los valores esperados son diferentes para ambas técnicas.

Tanto en visión lejana, como en cerca, ninguna de las gráficas de Bland y Altman del estudio de concordancia correspondientes a las medidas de la VFN ha mostrado tendencia a que la diferencia entre métodos aumente al aumentar el valor prismático,

es decir, el grado de acuerdo entre métodos es independiente del valor de los rangos de VFN medidos. Sin embargo, sí se puede observar cierta tendencia hacia el aumento de los CDCs al aumentar los rangos de VFP obtenidos (ver Figura 36 y Figura 37). Este comportamiento podría deberse al hecho de que en altos valores de vergencia fusional los pasos entre demanda prismática de la barra de prismas son mayores (por ejemplo 5Δ) comparada con los pequeños cambios de 1Δ en los diasporámetros, lo que podría generar mayor discrepancia entre métodos.

Como ocurría en el caso de la repetibilidad, tampoco se han encontrado estudios con los que comparar nuestros resultados. En el estudio de Ciuffreda *et al.*¹²⁵ con tres sujetos experimentados, comentado en el apartado de repetibilidad, encontraron que el promedio de las roturas y de los recobros BE obtenidos con barra de prismas fueron más altos (Ro: $39,1\Delta$; Re: 38Δ) que los obtenidos con los diasporámetros del foróptero (Ro: $32,3$; Re: $29,3\Delta$). Sin embargo, el promedio para los valores de rotura y recobro BI fueron similares en los dos entornos de medida (espacio libre y campo limitado del foróptero). Estos autores plantearon que los mayores rangos de vergencia medidos en el espacio libre respecto al foróptero podrían deberse a la influencia de la fusión periférica sobre la capacidad de vergencia fusional^{139, 143}. En nuestro caso, tanto en BE como en BI los valores promedio de emborronamiento y rotura obtenidos con diasporámetro fueron más elevados que los obtenidos con barra de prismas; sin embargo, los recobros fueron mayores al medirlos con barra de prismas. Este comportamiento pudo deberse a la mayor facilidad para alcanzar valores de punto de emborronamiento y rotura más elevados si la potencia prismática se incrementa poco a poco, como ocurre en el diasporámetro; frente a la barra de prismas donde la demanda de vergencia es asimétrica y los cambios se introducen a saltos, lo que puede ser más difícil para el paciente. Posiblemente, el efecto de la visión periférica comentado por Ciuffreda *et al.*¹²⁵ influya de forma más determinante cuando se trata

de retomar la visión simple tras la diplopía, lo que podría explicar el mayor recobro encontrado con barra de prismas frente al diasporámetro.

Correlación entre el PPC y la Rotura BE con diasporámetro o barra de prismas

Tanto el diasporámetro como la barra de prismas tienden a dar menores valores de amplitud máxima de convergencia en cerca que el PPC. Dada las diferencias metodológicas entre ambos tipos de pruebas, no sorprende demasiado la diferencia entre los valores de convergencia encontrados con las distintas pruebas. Respecto a los factores que podrían justificar la mayor convergencia en el PPC se puede decir que cabe esperar un mayor tiempo de respuesta al medir el PPC que al medir con prismas. Así, al aumentar la demanda de vergencias con prismas es frecuente que la separación de las imágenes al llegar al punto de rotura suceda de forma más brusca y marcada que en el PPC y al ser una situación más fácil de identificar por el paciente, es posible que se acorte el tiempo de respuesta. Por otra parte, la acomodación y la convergencia proximal podría también influir sobre los resultados. No obstante, no se espera que la influencia sea notoria pues recientemente Alder et al.¹²⁷ han visto que el punto de rotura en el PPC tiende a estar más alejado al utilizar la regla de cerca respecto a los estímulos presentados en el espacio libre, posiblemente debido a la distracción física que la regla produce en los sujetos. Tanimoto *et al.*¹⁴⁴ han mostrado que la dinámica de la respuesta inicial de vergencia puede estar influenciada por la disparidad periférica de modo que, aunque la regla en sí permanece estática, el hecho de ser vista en diplopía fisiológica también podría justificar el menor PPC al usarla. Además, dado que la regla utilizada para medir el PPC estuvo en contacto con la cara del paciente, podría haber una posible contribución asociada a la retroalimentación propioceptiva¹²⁷. Así, la retroalimentación propioceptiva podría indicar que la tarjeta está más cerca de lo que realmente está y generar un estímulo para cesar antes de

tiempo la convergencia y la acomodación, dando lugar a un PPC más alejado. Por todo lo dicho, si se usa regla de cerca para medir el PPC se debería comparar el resultado obtenido con los valores normales establecidos utilizando regla y un estímulo de fijación similar.

Conclusiones

Las dos técnicas analizadas para medir con prismas los rangos de vergencia fusional son bastante repetibles para la medida de VFN, siendo la repetibilidad reducida en la medida de VFP. La repetibilidad de la medida del PPC es buena.

Clínicamente hablando, la concordancia entre la medida con barra de prismas y con diasporámetro es moderada por ello el clínico debe ser cauto cuando compare los resultados de vergencia con barra de prismas con los obtenidos mediante los diasporámetros, especialmente en los rangos de vergencia BE. Además, la correlación entre los valores de convergencia medidos con estas dos técnicas respecto al PPC es baja. Por tanto, una vez que se elija una técnica inicial de medida, se deberá utilizar la misma en futuros exámenes de seguimiento del paciente.

Flexibilidad de vergencias

Introducción

La flexibilidad de vergencias (FV) se define como el número de ciclos por minuto que puede fusionarse un estímulo a través de una demanda prismática que se alterna entre base interna (BI) y base externa (BE). Con esta prueba se intenta evaluar la capacidad del sistema de vergencias para responder de forma rápida y con exactitud a cambios en la demanda de vergencia durante un periodo de tiempo, es decir, su capacidad para resistir a la fatiga. La flexibilidad de vergencias puede evaluarse en cerca o en lejos, en nuestro caso optamos por aplicar la prueba en visión próxima¹¹⁹.

Se han realizado diferentes investigaciones en las que se ha estudiado la potencia prismática y el tipo de prismas (flippers o prismas sueltos) más adecuados para la medida de flexibilidad de vergencias. Hay dos potencias prismáticas montadas en flippers que han sido especialmente estudiadas: 8Δ BI / 8Δ BE y 5Δ BI / 15Δ BE. En el estudio presente optamos por una potencia prismática del flipper de 3Δ BI y 12Δ BE que fue la potencia prismática para la que Gall *et al.*¹⁴⁵ encontraron una mayor utilidad clínica y buena repetibilidad en cerca.

En otro estudio, Gall *et al.*¹⁴⁶ analizaron la influencia del tipo de estímulo sobre la flexibilidad de vergencias y concluyeron que no es necesario utilizar tests especiales con controles de supresión (anaglifos o vectogramas) para obtener resultados válidos al medir FV, sino que una columna de letras es un buen test de fijación. Además, la presencia de una supresión durante la prueba se puede apreciar subjetivamente por el paciente por un desplazamiento lateral de la imagen única si se suprime el ojo sin prisma, o por no haber ningún cambio en la imagen si se suprime el ojo con prisma.

De forma objetiva el examinador podrá ver que los ojos del paciente realizan un movimiento versional sin respuesta de vergencias si se suprime el ojo con prisma o ningún movimiento si se suprime el ojo sin prisma.

Otro parámetro a tener en cuenta en la medida de la FV es la distancia de fijación pues influye sobre la cantidad de acomodación refleja y proximal y sobre la vergencia proximal estimulada. La mayoría de los estudios y opiniones están de acuerdo en la conveniencia de evaluar la FV a 40 cm donde entran en juego las complejas interacciones entre acomodación y vergencia¹⁴⁵.

El objetivo de este apartado de la investigación fue establecer la repetibilidad en la medida de la FV utilizando el método de Band y Altman.

Método

Durante las medidas de FV el sujeto llevaba puesta en gafa de pruebas su compensación de lejos. Como estímulo de fijación se usó una columna de letras de alto contraste correspondiente a una AV equivalente de Snellen de 0,8 situada en un atril a 40 cm de distancia¹⁴⁶. Se usó la luz de sala más un foco dirigido hacia el test (contraste \approx 80%). Para introducir la demanda prismática se utilizaron flippers cuya potencia fue de 3 Δ BI y 12 Δ BE. Se explicó al sujeto que las lentes que se interpondrían ante sus ojos le harían ver momentáneamente doble, es decir dos columnas de letras una al lado de la otra. El sujeto debía esforzarse por recuperar la fusión y por ver nítidas las letras tan pronto como le fuera posible, y decir “ya” una vez logrado. Se comenzó interponiendo los prismas base interna. El examinador se encargó de voltear el flipper cada vez que el paciente le avisaba. Se tuvo especial cuidado con mantener el eje longitudinal del flipper perpendicular a la línea media vertical del sujeto para no inducir demanda de vergencia vertical. A modo de

comprobación el examinador observaba los movimientos de refijación que hacían los ojos del sujeto al recuperar la fusión. El resultado de la prueba correspondió al número de ciclos realizados en 1 minuto (cpm). Un ciclo incluye el paso por BI y BE. En la primera tanda de medidas se dio al sujeto un periodo de prueba de 20 segundos antes de comenzar la medida.

1.18. Resultados

La Tabla 24 muestra los resultados de repetibilidad en la medida de la flexibilidad de vergencias. La prueba presenta una pobre repetibilidad con una DM significativa y un CDR elevado.

Tabla 24. Repetibilidad en la medida de flexibilidad de vergencias

	Media (cpm)	DE (cpm)	DM (F-I) (cpm)	p (test-t)	CDR (cpm)
FV	7,74	3,79	+1,96	p < 0,0001	±5,29
Clave: FV = flexibilidad de vergencias. DE = desviación estándar. DM = diferencia media. CDR = coeficiente de repetibilidad. F = final, I = inicial. cpm = ciclos por minuto.					

La Figura 39 muestra los resultados del estudio de repetibilidad para la medida de FV. En la gráfica se representan las diferencias vs. la media (Bland-Altman), el eje-y corresponde a la diferencia entre las dos medidas (Final – Inicial) y en el eje-x se refleja la media de las dos medidas para cada sujeto. La línea continua más gruesa corresponde a la diferencia media (DM), 1,96 cpm (p < 0,0001), que refleja un aumento en el número de ciclos en la segunda sesión respecto a los valores iniciales. Las líneas discontinuas corresponden a los límites de acuerdo al 95% (DM ± 5,29 cpm). La gráfica no muestra una tendencia a que la diferencia aumente con el rango de flexibilidad, es decir, la repetibilidad de la prueba no cambia con el número de ciclos alcanzado.

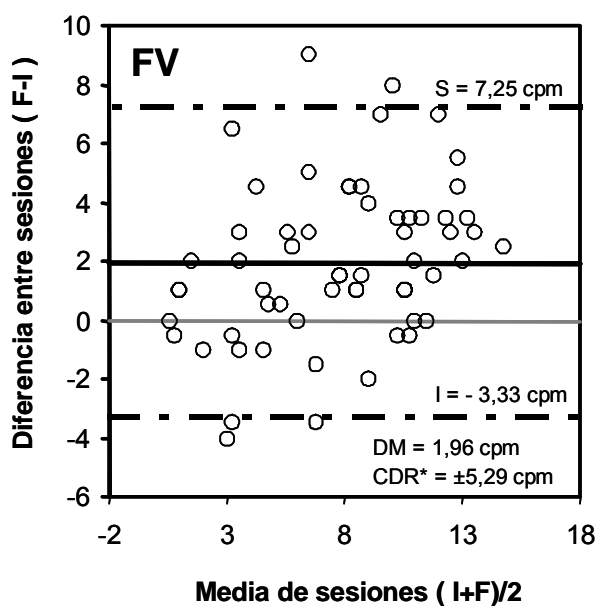


Figura 39. Diferencia media (DM), límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) y coeficiente de repetibilidad (CDR) para la medida de flexibilidad de vergencias (FV) con flippers de 3Δ BI / 12Δ BE..

1.19. Discusión

Para la flexibilidad de vergencias se obtuvo una media de 7,74 cpm y una DE de 3,79 cpm. Este valor medio fue notoriamente menor al establecido por Gall *et al.*¹⁴⁵ con un flipper de la misma potencia (13,5 cpm, DE = 4,0). Es posible que las condiciones de laboratorio bajo las que se tomaron las medidas en el presente estudio pudieran haber distraído a los sujetos justificando, en parte, la baja repetibilidad de la prueba y a los bajos valores medios encontrados. Otros estudios reflejan valores menores que los de Gall *et al.*¹⁴⁵, aunque la demanda prismática BI utilizada fue superior (8Δ BI y 12Δ BE). Por ejemplo, Hoffman y Rouse¹⁴⁷ vieron que en sujetos normales se esperan unos resultados entre 8 y 10 cpm y Waugh¹⁴⁸ encontró unos valores muy parecidos a los del presente estudio (7,7 cpm, DE = 3,2).

Hay pocos trabajos en los que se estudie la repetibilidad de la medida de flexibilidad de vergencias en el rango de edad de nuestra población, y en ninguno de ellos se hace un estudio Bland-Altman, sino uno de correlación. Se han encontrado los siguientes resultados¹⁴⁵, con 3Δ BI / 12Δ BE en cerca, $r = 0,85$; con 8Δ BI / 12Δ BE en cerca, $r = 0,75$. A partir de las medidas de este estudio se obtuvo también una alta correlación con $r = 0,79$ (ver Figura 40). La línea de puntos corresponde a la recta de referencia $x = y$, mientras que la línea continua corresponde al ajuste lineal por mínimos cuadrados cuya ecuación y valor r se muestra en la gráfica. La alta correlación encontrada entre la medida inicial y final parece indicar una buena repetibilidad de la prueba de FV, lo que está en contradicción con la baja repetibilidad que refleja el estudio Bland y Altman. Siendo éste un buen ejemplo de la conveniencia de utilizar los métodos estadísticos adecuados para este tipo de análisis (ver Anexo V).

La DM encontrada (1,96 cpm) en el estudio de Bland y Altman y el punto de corte con el eje de ordenadas de la recta resultante del ajuste lineal ($y_0 = 2,08$ cpm) nos hace pensar que el efecto aprendizaje influye notoriamente en la repetibilidad de la medida de FV.

Aunque hay pocas razones para pensar que la dinámica del sistema de vergencias se vea comprometido con la edad, aunque sí se ha visto que hay una tendencia al aumento de la exoforia con los años^{149, 150}, probablemente por la disminución de la convergencia acomodativa a través de las adiciones positivas. Pellizzer y Siderov⁶

estudiaron la repetibilidad de la FV en 21 sujetos asintomáticos de 36 a 77 años (media = 58,7, DE = 10,6) que lograron fusionar ambas direcciones del flipper en las 4 sesiones de medida llevadas a cabo. Usaron un flipper de 8Δ BI / 12Δ BE colocado delante de la propia compensación del sujeto de visión próxima. La media de FV entre todos los sujetos y sesiones de medida fue 6,7 cpm (DE = 2,1) que comparándola con los resultados de estudios llevados a cabo con poblaciones más jóvenes^{147, 148} y con el estudio presente, parece confirmar que la FV no disminuye significativamente con la edad. Los límites de acuerdo al 95% entre la sesión 3 y la 4 fue $CDR = \pm 8,9$ cpm, y entre la 1 y la 4 fue $CDR = \pm 5,2$ cpm. Aunque un análisis ANOVA reveló que no había diferencias significativas entre sesiones de medida ($p > 0,05$).

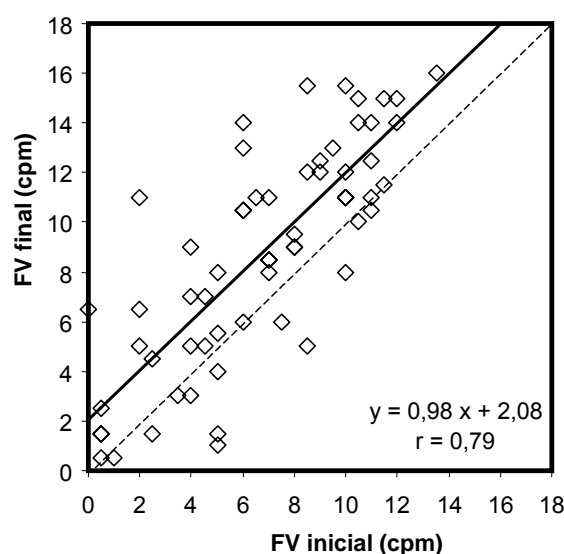


Figura 40. Correlación entre las medidas de flexibilidad de vergencias inicial y final.

En el estudio de Pellicer y Siderov⁶ también se comparó el número de ciclos conseguido en los primeros 30 s y en los siguientes 30 s y los resultados revelaron que no había diferencias significativas entre ambos intervalos, pudiéndose concluir que en sujetos asintomáticos al medir la FV durante un periodo de 1 min la fatiga no

afecta de forma apreciable a los resultados y que la diferencia entre sus CDR se debiera simplemente a la alta variabilidad de los resultados de la prueba.

1.20. Conclusiones

La buena correlación entre medidas repetidas la FV hace pensar que se trata de una de las pruebas clínicas de visión binocular más repetibles. Sin embargo, el estudio de Bland y Altman mostró que en realidad es una prueba con una pobre repetibilidad y que diferencias de hasta 5,29 cpm por encima o por debajo de medidas previas del paciente pueden deberse a errores propios de la prueba y no a un cambio real en las capacidades del sujeto evaluado.

Estereoagudeza

Introducción

Hay muchas claves para la percepción de la posición relativa de los objetos en el espacio: superposición, perspectiva, sombras, paralaje, tamaño, acomodación, convergencia y estereopsis¹³. La visión estereoscópica es una característica fundamental de la visión binocular humana. La separación horizontal entre los dos ojos hace que se reciban imágenes ligeramente dispares de los objetos y esta disparidad es utilizada para tener una percepción de su profundidad relativa. La estereoagudeza o estereopsis corresponde al mejor nivel posible de las funciones binoculares. No obstante, los beneficios funcionales de una buena estereopsis se han despreciado frecuentemente, quizá porque el impacto real de la pérdida de estereoagudeza en la vida diaria puede ser sólo mínimo^{151, 152}.

La constatación de la existencia del nivel de estereoagudeza esperado para la edad es un dato muy interesante para los profesionales del cuidado de la visión pues indica un desarrollo normal de las funciones sensoriales y motoras¹⁵³. Esto se debe a que para tener una buena estereoagudeza es necesario que la agudeza visual sea igual en los dos ojos, que los ojos estén correctamente alineados y que el paciente haya tenido experiencia previa de binocularidad. Por ello los tests para medir la estereoagudeza en visión próxima se utilizan comúnmente en la práctica clínica para identificar anomalías en las funciones binoculares tales como heterotropía, ambliopía y otras alteraciones de la visión binocular^{154, 155}.

Actualmente hay diversos tests disponibles en el mercado para evaluar la estereoagudeza y pueden dividirse en tres categorías:

- (1) Estereogramas locales o tests de contornos, en los que se crea profundidad introduciendo una disparidad horizontal entre los límites de los dibujos vistos por cada ojo. La mayor desventaja de estos tests es la presencia de pistas monoculares de profundidad que pueden ser usadas por un observador astuto para elegir la opción correcta en ausencia de estereopsis¹⁵⁶.
- (2) Estereogramas de puntos al azar o test de estereopsis global. Fueron introducidos por Julesz¹⁵⁷ y en ellos la profundidad se crea mediante desplazamiento horizontal de un grupo de elementos aleatorios. Su principal ventaja es que no muestran pistas monoculares.
- (3) Tests de profundidad real. En ellos la profundidad real se crea mediante una separación real por lo que no es necesario equipar al paciente con filtros disociantes, sino que las imágenes son vistas por ambos ojos. También presentan pistas monoculares como el paralaje y las sombras¹⁵⁴.

El test tradicionalmente más usado ha sido el test de Titmus, pero con la introducción de estereogramas de puntos al azar, han aparecido nuevos tests, por ejemplo, el test Random-dot-E, el test TNO, el test de Randot y el test de Frisby entre otros. En general, todos ellos son fáciles de administrar aunque cada uno tiene sus propias características, haciendo que haya grandes variaciones de precisión entre ellos¹⁵⁸. Hay otros factores, además de la facilidad de comprensión, que son muy difíciles de controlar; en particular, una mala cooperación afecta a la precisión de la prueba^{159, 160}. Un buen test debe ser cuidadosamente diseñado para asegurar que sólo evalúa la estereopsis. En general, los estereotests de puntos al azar que miden la estereopsis global son más efectivos para detectar estrabismo, ambliopía, baja AV o anisometropía que los estereotests de contornos que miden estereopsis local. Esto se debe a que, como ya se ha comentado, los estereotests de contornos presentan pistas

monoculares, contornos o desplazamientos laterales que reducen su validez y que están ausentes en los estereotests de puntos al azar^{14, 161}.

Entre los tests que presentan pistas monoculares se han citado el de Titmus, el de Randot y el de Frisby¹⁵⁵. El test de Titmus y el de Randot (versión 2), contienen pistas monoculares que hacen posible llegar hasta cierto nivel sin tener realmente ese grado de estereopsis. Cuando se ve el test bajo condiciones que impiden presentar estereoagudeza o cuando es visto por pacientes con anomalías conocidas de visión binocular, en el test de Titmus se puede identificar correctamente el anillo de Wirt hasta el tercer o cuarto grupo, y hasta el primer o segundo grupo de círculos en el test de Randot (v 2)^{156, 162-164}. Como resultado de ello se puede asignar a muchos pacientes un nivel de estereoagudeza erróneo, concluyéndose por ello que han recuperado la visión binocular cuando, o no tienen estereopsis, o sólo presentan una estereopsis gruesa.

Los estereotests de profundidad real, como el test de Frisby, tienen la ventaja de no necesitar gafas de disociación y por lo tanto son menos artificiales. Aunque se ha demostrado que el test de Frisby también presenta pistas monoculares debido a que está compuesto de dibujos (triángulos) relativamente grandes (hasta 1,4 grados de ángulo visual a 40 cm)¹⁶⁵. Sin embargo, varios estudios llevados a cabo, tanto en niños como en adultos, han encontrado que el test de Frisby no puede pasarse monocularmente cuando se minimizan las pistas debidas al paralaje¹⁵⁵.

Por otro lado, se ha comprobado que la disociación con filtros rojo-verde puede mejorar o degradar artificialmente la visión binocular. Por ejemplo, varios pacientes con estereoagudeza anormalmente alta a pesar de tener anisometropía o ambliopía anisométrica, mostraron esos niveles de estereoagudeza cuando el ojo ambliope o el más amétrope veía la imagen verde del test TNO (filtro rojo). Esto parece deberse a

que el mayor contraste de la mitad verde respecto a la mitad roja del dibujo del TNO (con una diferencia 2:1 o mayor) compensa la reducción de sensibilidad al contraste del ojo ambliope, mejorando la estereoagudeza¹⁶⁶.

Como consecuencia de las diferencias entre los tests y de las limitaciones de cada uno de ellos, resulta muy difícil comparar los resultados de estereoagudeza medidos con distintos tests de estereoagudeza¹⁶⁷ y establecer unos criterios generales de normalidad. Los resultados dependen mucho del método de medida (global o local), de las condiciones bajo las que se aplica el test (del nivel de iluminación y contraste, de si se realiza la prueba en el espacio libre o a través de un instrumento, de si es necesario interponer filtros anaglifos o polarizados o no, de lo cansado que esté el paciente al realizar la prueba, etc.) y del criterio de umbral elegido^{58, 168}.

El objetivo de este estudio fue evaluar la repetibilidad intra-examinador de algunos de los tests que más frecuentemente se utilizan en la práctica clínica habitual para obtener el umbral de estereoagudeza (Frisby, TNO, Randot y Titmus), así como establecer el nivel de concordancia entre los cuatro tests en una población de adultos jóvenes con visión binocular normal.

Método

Se midió la estereoagudeza a todos los sujetos mediante cuatro pruebas distintas. La Tabla 25 describe las partes de que consta cada test y su grado de disparidad. Todas las pruebas fueron administradas por el mismo examinador y cada una de ellas se aplicó exactamente de la misma forma en todos los sujetos. A la hora de aplicar los tests se siguieron las indicaciones que aparecían en las instrucciones de cada prueba. En todos los participantes se midió en primer lugar con el test de Frisby pues es el único de los tests que no precisa ningún sistema de disociación, mientras que el orden

de las restantes tres pruebas (TNO, Randot y Titmus) se estableció de forma aleatoria mediante extracción de bolas numeradas.

Tabla 25. Descripción de los tests de estereoagudeza

Test	Características del test	Grado de disparidad (seg arc)	Distancia al test
Frisby	<ul style="list-style-type: none"> - 3 láminas de diferente espesor (elementos al azar)* - No necesita filtros (test de profundidad real) 	150	30 cm
		85	40 cm
		55	50 cm
		40	60 cm
		30	70 cm
		20	80 cm
TNO	<ul style="list-style-type: none"> - 3 láminas con 6 grados de disparidad (diseño de puntos al azar) - Precisa gafas anaglifas (rojo-verde) 	480	Siempre se mide a 40 cm
		240	
		120	
		60	
		30	
Randot	<ul style="list-style-type: none"> - 10 grupos de círculos (diseño de contornos con fondo de puntos al azar) - Precisa gafas polarizadas 	15	Siempre se mide a 40 cm
		400	
		200	
		140	
		100	
		70	
		50	
		40	
		30	
25			
Titmus	<ul style="list-style-type: none"> - 9 grupos de círculos (diseño de contornos) - Precisa gafas polarizadas 	20	Siempre se mide a 40 cm
		800	
		400	
		200	
		140	
		100	
		80	
		60	
50			
40			

* Sólo se reflejan los valores de disparidad para la lámina de 1,5 mm de espesor

A continuación se comentan algunas de las características de cada test y se describe con detalle el procedimiento de medida seguido con cada uno de ellos:

FRISBY

El test de Frisby (Clement Clarke, Harlow, UK) se compone de tres láminas de prexiglas transparente de diferentes grosores (1,5 mm, 3 mm y 6 mm). Cada lámina contiene cuatro cuadrados en cuyo interior aparecen pintadas pequeñas figuras geométricas distribuidas al azar. Uno de los cuatro cuadrados tiene una parte de sus pequeñas figuras, formando un círculo, pintadas en la otra cara de la lámina (ver Figura 41). Los otros tres cuadrados están pintados en su totalidad en la misma cara de la lámina. Este diseño permite evaluar la capacidad de los sujetos para percibir en profundidad. En concreto, se evaluó la capacidad para detectar la disparidad cruzada^{††}, comprobando si el sujeto podía identificar en cuál de los cuatro cuadrados había un círculo de figuras de 3 cm de diámetro que estaba por delante del resto de las figuras.

Se informó a cada sujeto que su tarea consistía en identificar cuál de los cuatro cuadrados tenía inscrito un círculo de figuras que flotaba por delante y se comprobó mediante la lámina más gruesa que entendían su tarea. A continuación se pasó a identificar cuál era el umbral de estereoagudeza, buscando la combinación de menor grosor de lámina y máxima distancia a la que el sujeto era capaz de identificar correctamente tres de cuatro presentaciones. La lámina se retiraba brevemente de la vista del sujeto mientras se cambiaba su orientación.

^{††} De modo similar se podría evaluar la capacidad para percibir la disparidad descruzada, comprobando si el sujeto es capaz de identificar en cual de los cuatro cuadrados hay un círculo de figuras que está por detrás del resto de las figuras. En nuestro estudio para no alargar mucho las medidas nos limitamos a evaluar la percepción de la disparidad cruzada tal y como se recomienda en aquellos casos donde sólo se evalúe una dirección de disparidad (Saladin JJ. In Benjamin WJ (ed.), *Clinical Refraction*. Philadelphia: Saunders Co; 2006: 899-960)

El test se comenzó realizando a 40 cm, lo que permitió conseguir una estereoagudeza máxima de 85 seg arc para una distancia interpupilar de 60 mm. Si el sujeto superaba adecuadamente la prueba a esta distancia se pasaba a presentar la lámina más fina a mayores distancias en pasos de 10 cm. En todos los casos las láminas se mostraron perpendicularmente a la línea de mirada contra un fondo blanco, se evitó que hubiera reflejos sobre la lámina presentada y no se permitió al sujeto que moviera la cabeza o la lámina; de este modo se intentó evitar que obtuviera pistas debidas a efectos de paralaje¹⁵⁵.



Figura 41. Estereoagudeza con el test de Frisby

El test incluye una tabla que correlaciona la combinación espesor-distancia con la medida de umbral de disparidad correspondiente. En general, se espera que un sujeto adulto con visión binocular normal muestre un umbral de 20 seg arc con este test.

TNO

El test TNO (Lameris Instrumenten, Groenekan, ND) está formado por un conjunto de estereogramas donde las imágenes vistas por cada ojo están superpuestas y han sido impresas en colores complementarios. Para la observación de estos estereogramas se equipa al sujeto con gafas anaglíficas (rojo-verde). El diseño de estos estereogramas está basado en el principio de puntos al azar introducido por Julesz¹⁵⁷. En este tipo de

láminas cuando se observan por separado las imágenes de uno u otro ojo no se obtiene ninguna sensación de profundidad relativa; esto es debido a que no contienen ninguna pista monocular.



Figura 42. Estereoagudeza con el test TNO

El test se compone de siete láminas: las primeras cuatro láminas se suelen usar para *screening*, mientras que las tres últimas presentan figuras que cubren un rango de estereopsis que va de 480 a 15 seg arc y que permiten establecer el grado de estereoagudeza del sujeto. Estas tres últimas láminas fueron las que se mostraron a los sujetos que participaron en nuestro estudio (ver Figura 42). Para cada nivel de estereoagudeza evaluado el test contiene dos discos de 5,6 cm de diámetro a los que les falta un sector circular orientado en determinada dirección (como si fuera un pastel al que le falta un trozo). El cometido del sujeto consistió en determinar la orientación (arriba, abajo, derecha o izquierda) del sector que faltaba. El resultado de la prueba correspondía al último nivel en el que el sujeto respondió correctamente sobre los dos componentes de dicho nivel. En aquellos casos en los que el sujeto respondió de forma incorrecta a una de las dos figuras se le permitió dar una segunda respuesta sobre la figura fallada; si en este segundo intento la respuesta de nuevo era incorrecta se anotó como resultado el nivel anterior. Si en el segundo intento la respuesta era correcta se daba por superado el nivel y se le pidió que intentara determinar la

orientación de la abertura correspondiente al siguiente nivel de estereoagudeza. Para este test se considera clínicamente aceptable un resultado de 60 seg de arco¹⁶⁷.

No se estableció un tiempo máximo para la realización de la prueba, incluso en aquellos sujetos que dudaban a la hora de dar una contestación se les informó de que podían tomarse el tiempo necesario para averiguar la respuesta.

Para realizar la prueba se equipó a los sujetos con su compensación de lejos sobre la que se colocaron las gafas anaglifas. El test se situó en un atril para poder controlar que todos los sujetos lo observaran bajo un ángulo de 45° respecto al plano facial y a una distancia de 40 cm. Como iluminación, además de la luz de sala se situó una lámpara adicional por encima del hombro del sujeto para lograr una iluminación correcta del test sin sombras.

RANDOT

El test de Randot versión 2 (Stereo Optical Company, Chicago, IL) es un test de estereopsis muy utilizado de estereoagudeza que tiene una parte de puntos al azar correspondiente a dos niveles de disparidad grandes (500 y 250 seg arc) y otra parte híbrida de estímulos de contornos sobre un fondo de puntos al azar con la que se cubre un amplio rango de disparidades (de 400 a 20 seg arc) (ver Figura 43). Durante la aplicación del test el sujeto llevaba su compensación de lejos y gafas polarizadas (ejes a 45° y 135°). El test se situó en un atril para poder controlar que todos los sujetos lo observaran bajo un ángulo de 45° respecto al plano facial y a una distancia de 40 cm. El test se mantuvo uniformemente iluminado y sin reflejos que pudieran dificultar la medida.

Como la muestra de sujetos la componían sujetos adultos con visión binocular normal, la medida de estereoagudeza se realizó directamente con los anillos de Wirt. Se

informó al sujeto que su cometido en la prueba consistía en identificar, para cada grupo de tres anillos, cuál era el anillo que parecía flotar por encima de los demás^β (debía decirnos: “derecha”, “centro” o “izquierda” en función de la posición que adoptara el anillo seleccionado). Se comenzó la prueba por el primer grupo de anillos situado arriba a la izquierda y se continuó con el grupo número dos, tres, etc. Se anotó como resultado de la prueba la estereoagudeza correspondiente al último grupo de anillos correctamente identificado antes de 2 fallos consecutivos. Si el sujeto fallaba en un grupo de anillos, pero acertaba el siguiente, se le daba una segunda oportunidad para identificar correctamente el anillo anteriormente fallado. Se espera que un sujeto con visión binocular normal alcance 20 seg arc con esta prueba.



Figura 43. Estereoagudeza con el test de Randot

TITMUS

El estereotest de Titmus (Stereo Optical Company, Chicago, IL) se compone de tres partes: La mosca, los animales y los anillos de Wirt (ver Figura 44). Durante la realización del test se tuvieron en cuenta las mismas condiciones que en el test de Randot y también se comenzó midiendo directamente con los puntos de Wirt.

^β Uno de los anillos en cada grupo contiene un cierto grado de disparidad cruzada, de modo que cuando se observa binocularmente parece flotar por encima de los demás anillos del grupo.

Se informó al sujeto que su cometido en la prueba consistía en identificar para cada grupo de cuatro anillos, cuál era el anillo que parecía flotar por delante de los demás (debía decirnos “arriba”, “abajo”, “derecha” o “izquierda” en función de la posición que adoptara el anillo seleccionado). Se comenzó la prueba por el primer grupo de anillos situado arriba a la izquierda y se continuó según el orden de lectura. Se anotó como resultado de la prueba la estereoagudeza correspondiente al último grupo de anillos correctamente identificado antes de 2 fallos consecutivos. Si el sujeto fallaba en un grupo de anillos, pero acertaba el siguiente, se le daba una segunda oportunidad para identificar correctamente el anillo anteriormente fallado.



Figura 44. Estereoagudeza con el test de Titmus

El rango de disparidades de los anillos de Wirt va de 800 a 40 seg arc cuando se aplica la prueba a 40 cm. En un sujeto con visión binocular normal se espera que la estereoagudeza con este test sea 40 seg arc, que es el máximo nivel evaluado a 40 cm.

Resultados

REPETIBILIDAD

La Tabla 26 muestra los resultados del estudio sobre la repetibilidad de las cuatro técnicas de medida de estereoagudeza seleccionadas. La mejor repetibilidad correspondió al test de Randot pues presentó la menor diferencia media entre las dos sesiones de medida (-1,23; $p=0,5$) y el menor intervalo de acuerdo (± 25 seg arco), reflejando la menor variación entre las dos sesiones de medida.

Tabla 26. Repetibilidad en la medida de la estereoagudeza en visión próxima

	Media (seg arco)	DE (seg arco)	DM (F-I) (seg arco)	p(test-t)	CDR (seg arco)
Frisby	24,10	9,87	-2,30	$p = 0,2$	$\pm 35,00^*$
TNO	66,93	66,42	-3,52	$p = 0,4$	$\pm 60,00^*$
Randot	31,52	13,30	-1,23	$p = 0,5$	$\pm 25,00^*$
Titmus	46,43	19,70	2,38	$p = 0,3$	$\pm 40,00^*$

Clave: DE = desviación estándar. DM = diferencia media. CDR = coeficiente de repetibilidad, seg arco = segundos de arco. F = final, I = inicial
 * En los casos donde CDR aparece con un asterisco, se ha sustituido éste por el valor del percentil 95 del valor absoluto de la diferencia

La Figura 45 muestra las gráficas de Bland y Altman para las cuatro técnicas de medida de la estereoagudeza. El eje-y corresponde a la diferencia con signos entre las medidas de ambas sesiones (Final – Inicial) y el eje-x corresponde a la media de las dos medidas. Estas gráficas facilitan la comparación del grado de repetibilidad de las distintos tests, pues cuanto menor sea la diferencia media entre sesiones de medida (DM) y más estrecho sea el intervalo de acuerdo al 95%, más repetible será una técnica. Así, dado que la escala del eje-y es la misma para todas las gráficas es fácil comprobar que la peor repetibilidad correspondió al test TNO. Todas las gráficas parecen mostrar una cierta tendencia a que la diferencia aumente al aumentar la disparidad (especialmente a partir de unos 50 seg arco), es decir, la repetibilidad de las

pruebas empeora al disminuir el grado de estereoagudeza (mayor grado de disparidad).

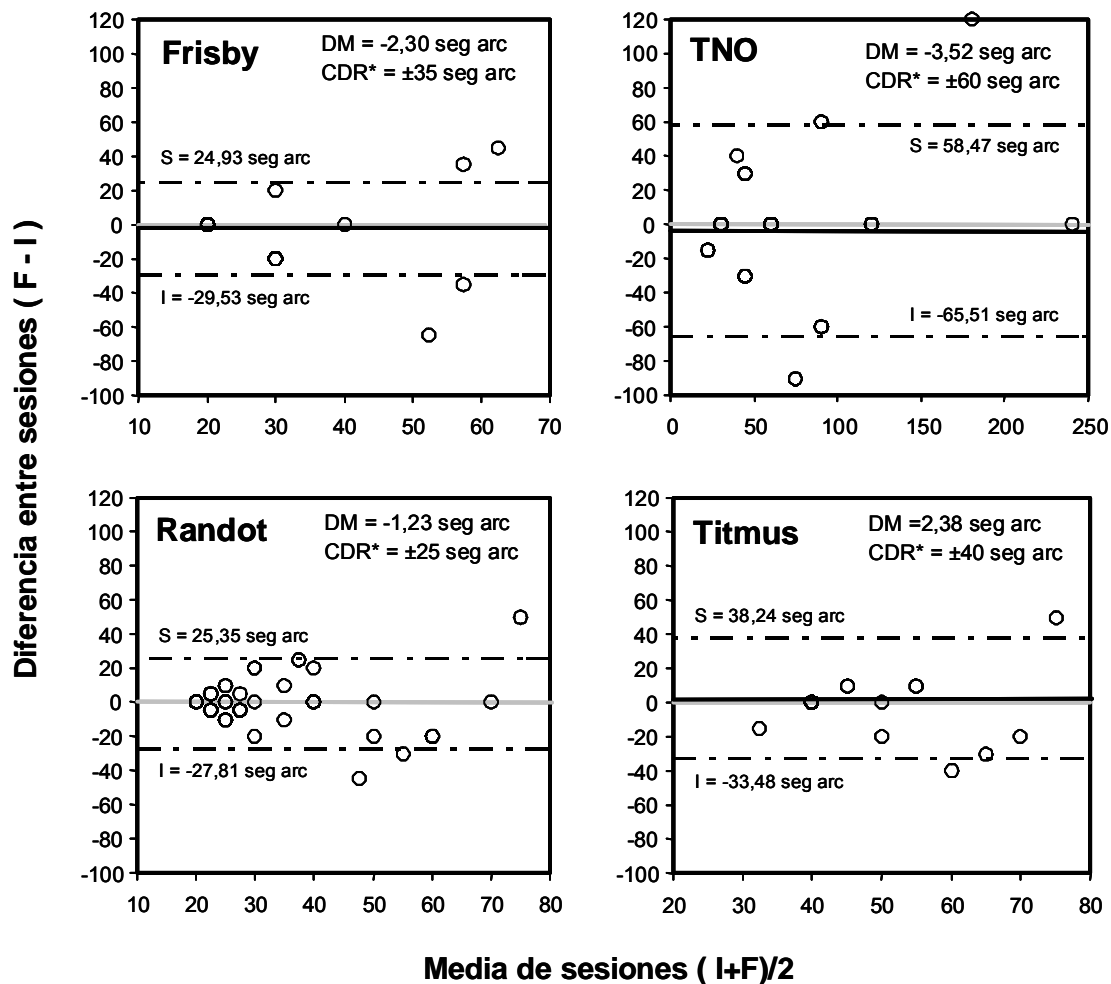


Figura 45. Diferencia media (DM) y límites del intervalo de acuerdo al 95% de Bland y Altman, superior (S) e inferior (I) para los diferentes tests de medida de la estereoagudeza en cerca. En todos los casos la distribución de las diferencias no fue normal y se tomó como coeficiente de repetibilidad el percentil 95 del valor absoluto de las diferencias (CDR*).

En los casos en los que la repetibilidad cambia en función del valor de la variable medida Bland y Altman²⁸ sugieren separar el rango de valores en tramos y evaluar la repetibilidad en cada uno de ellos. Por ello, separamos la muestra en dos partes en cada test según tuvieran un valor medio de estereoagudeza entre sesiones < 50 seg

arco ó ≥ 50 seg arco para la prueba correspondiente y procedimos a estudiar la repetibilidad en cada tramo. La Tabla 27 refleja los resultados obtenidos de este análisis y aunque el número de sujetos en alguno de los subgrupos fue reducido, sí se puede ver que, efectivamente, para todos los estereotests la repetibilidad fue notoriamente mejor en el tramo correspondiente a una mejor estereoagudeza (< 50 seg arco) y peor en el tramo de peor estereoagudeza (≥ 50 seg arco) y que además, la repetibilidad obtenida cuando se tomaba todo el rango de valores se situaba entre ambas (ver Tabla 26).

Tabla 27. Repetibilidad en la medida de estereoagudeza en visión próxima por tramos

	Rango (seg arco)	N	DM (F-I) (seg arco)	CDR (seg arco)
Frisby	< 50	57	-2,11	$\pm 20,00^*$
	≥ 50	4	-5,00	$\pm 62,00^*$
TNO	< 50	31	-3,39	$\pm 30,00^*$
	≥ 50	30	-4,00	$\pm 90,00^*$
Randot	< 50	53	-0,28	$\pm 22,00^*$
	≥ 50	8	-7,50	$\pm 43,00^*$
Titmus	< 50	50	0,10	$\pm 5,50^*$
	≥ 50	11	12,73	$\pm 80,00^*$

Clave: DE = desviación estándar. DM = diferencia media. CDR = coeficiente de repetibilidad, seg arco = segundos de arco. F = final, I = inicial

* En los casos donde CDR aparece con un asterisco, se ha sustituido éste por el valor del percentil 95 del valor absoluto de la diferencia

Los individuos con visión binocular normal pueden tener un umbral de estereoagudeza mejor de 10 seg arco¹³. Dado que la menor disparidad disponible en los tests clínicos es considerablemente mayor (ver Tabla 25), muchos de los sujetos pueden identificar la menor disparidad ofrecida por el tests. En el test de Frisby y en el de Titmus los datos de estereoagudeza aparecieron sesgados ya que se produjo este efecto techo

debido a que la mayoría de los sujetos (todos menos trece en ambas pruebas) alcanzaron el nivel de disparidad más bajo medible con el test según la metodología descrita (Frisby: 20 seg arc; Titmus: 40 seg arc) en ambas tandas de medidas^{§§}. El efecto techo que hizo que muchos sujetos presentaran una diferencia “cero” entre sesiones, unido a que los niveles de estereoagudeza estudiados en los distintos tests no están equiespaciados (ver Tabla 25) contribuyó a que la distribución de las diferencias no fuera normal en ninguno de los cuatro tests estudiados.

CONCORDANCIA

La Tabla 28 muestra los resultados del estudio de concordancia entre los distintos tests de medida de estereoagudeza. Las diferencias medias (DM) negativas asociadas al test de Frisby indican que este test fue el que tendió a arrojar menores valores como promedio (mejor estereopsis). La peor concordancia lo mostró el test TNO con las demás pruebas tal y como reflejan los amplios intervalos de acuerdo, llegándose a alcanzar ± 100 seg arc entre TNO y Frisby. La mayor concordancia apareció entre el test de Frisby y el de Randot, con un CDC = ± 30 seg arc y la menor DM = -7,42 seg arc.

La Figura 46 muestra las gráficas de Bland y Altman para el estudio de la concordancia entre métodos que miden la estereoagudeza, el eje-x corresponde a la media entre métodos y el eje-y a su diferencia. Las líneas horizontales representadas en estas gráficas indican la diferencia media (DM) y los límites de acuerdo al 95% entre parejas de métodos, $DM \pm 1,96 \cdot DE$ de las diferencias. Comparando estas

^{§§} Este efecto techo ya ha sido referenciado por otros autores tal y como comenta Hall en su artículo titulado: “The relationship between clinical stereotests”(Ophthal Physiol Opt 1982; 2:133-143). Hall propone aumentar la distancia a la que se aplican los tests para aumentar el rango de disparidad estudiado, sin embargo, esta estrategia no fue una solución totalmente práctica. Aplicando el test de Titmus a 80 cm, todavía cerca del 80% de los sujetos alcanzaba el máximo nivel de estereoagudeza (10 seg arc). Por otra parte, en el test de Frisby resulta ser un inconveniente práctico importante la incorporación nuevas distancias de observación más alejadas.

gráficas, teniendo en cuenta que la escala para el eje-y no es la misma para todos los casos, es fácil apreciar que hay notorias diferencias en el grado de acuerdo existente entre los distintos métodos. Fijándonos en las gráficas que comparan el test TNO con los demás tests estudiados, se observa una marcada correlación entre la media y la diferencia entre métodos, de modo que la diferencia entre el test TNO con los demás métodos aumenta al aumentar el grado de disparidad necesario para ser detectado (peor estereoagudeza). Dado este comportamiento, procedimos a estudiar la concordancia del TNO con los demás tests por tramos. Para ello separamos la muestra de sujetos en dos partes según tuvieran un valor medio de estereoagudeza < 50 seg arco ó ≥ 50 seg arco para la pareja de tests correspondiente.

Tabla 28. Concordancia entre pruebas que miden la estereoagudeza en visión próxima

	DM (seg arc)	p (test-t)	CDC (seg arc)
Frisby – TNO	-42,83 (TNO > Frisby)	p < 0,0001	±100*
Frisby – Randot	-7,42 (Randot > Frisby)	p < 00001	±30*
Frisby – Titmus	-22,34 (Titmus > Frisby)	p < 0,0001	±45*
TNO – Randot	35,41 (TNO>Randot)	p < 0,0001	±95*
TNO – Titmus	20,49 (TNO > Titmus)	p = 0,003	±80*
Randot - Titmus	-14,92 (Titmus > Randot)	p < 0,0001	±30*

Clave: DM = diferencia media. CDC = coeficiente de concordancia, seg arco = segundos de arco.
 * En los casos donde CDC aparece con un asterisco, se ha sustituido éste por el valor del percentil 95 del valor absoluto de la diferencia

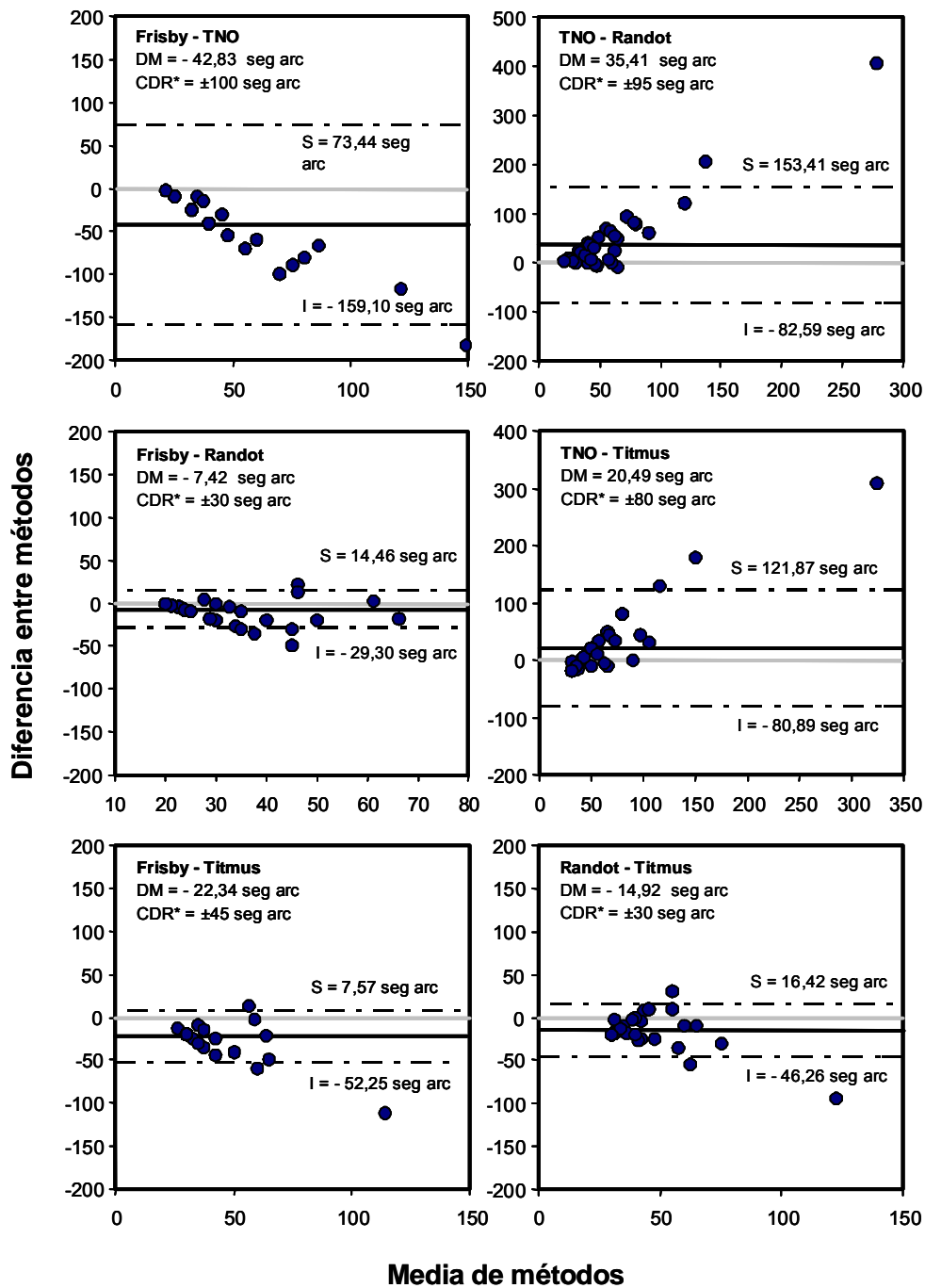


Figura 46. Gráficas de Bland y Altman para el estudio de la concordancia entre métodos que miden la estereoagudeza en visión próxima. La línea continua representa la diferencia media (DM) entre las medidas obtenidas con los distintos métodos. Las líneas discontinuas indican los límites superior (S) e inferior (I) del intervalo de acuerdo al 95% ($DM \pm 1,96 DE$). En todos los casos la distribución de las diferencias no fue normal y se tomó como coeficiente de concordancia el percentil 95 del valor absoluto de las diferencias (CDC*).

La Tabla 29 refleja los resultados obtenidos de este análisis se puede ver que en los tres casos la concordancia fue notoriamente mejor en el tramo correspondiente a una mejor estereoagudeza (< 50 seg arc) y peor en el tramo de peor estereoagudeza (≥ 50 seg arc) y que además esta última mejoraba respecto a la concordancia obtenida cuando se tomaba todo el rango de valores (ver Tabla 28).

Tabla 29. Concordancia del test TNO con los demás estereotests por tramos

	Rango (seg arc)	N	DM (seg arc)	CDC (seg arc)
Frisby - TNO	< 50	35	-4,42	$\pm 25,30^*$
	≥ 50	26	-30,13	$\pm 92,13^*$
TNO- Randot	< 50	31	0,92	$\pm 9,88^*$
	≥ 50	30	14,69	$\pm 41,92^*$
TNO - Titmus	< 50	30	1,62	$\pm 14,83^*$
	≥ 50	31	21,79	$\pm 61,31^*$

Clave: DE = desviación estándar. DM = diferencia media. CDR = coeficiente de repetibilidad, seg arco = segundos de arco. F = final, I = inicial

* En los casos donde CDC aparece con un asterisco, se ha sustituido éste por el valor del percentil 95 del valor absoluto de la diferencia

ANOVA

Los resultados de un ANOVA de medidas repetidas con dos factores (método y sesión) mostraron diferencias significativas en función del método aplicado ($F_{3,180}=25,05$; $p<0,0001$). Sin embargo, no se encontraron diferencias significativas entre las dos sesiones de medida ($F_{1,60}=0,69$; $p=0,4$), ni hubo efectos de interacción entre ambos factores ($F_{3,180}=0,96$; $p=0,4$). Un análisis *post-hoc* utilizando la técnica de Scheffé indicó que había diferencias significativas entre las medias del test TNO respecto al test de Frisby y respecto al test de Randot.

Discusión

REPETIBILIDAD

Es importante conocer el CDR al medir la estereoagudeza por métodos subjetivos pues establece el mínimo cambio en la estereoagudeza que claramente excede el margen de error inherente a la técnica para cada rango de edad. El menor margen de error, menor CDR, se ha obtenido para el test de Randot.

Una diferencia significativamente distinta de cero entre dos medidas tomadas siguiendo la misma metodología podría indicar que la primera medida influye sobre la segunda (por ejemplo, debido a la fatiga o al aprendizaje). En el presente trabajo las diferencias medias encontradas (ver Tabla 26) no fueron significativamente distintas de cero tanto desde el punto de vista estadístico como clínico, por lo que se puede concluir que esos efectos comentados no han influido notoriamente en el resultado de los tests.

La repetibilidad de una prueba puede estar influenciada por diferentes factores¹⁶⁹, incluida la edad y el nivel de la función que está siendo medida por la prueba. Algunos tests de estereoagudeza pueden tener un buen acuerdo test-retest cuando se mide a sujetos con función sensorial binocular normal, pero tener peor acuerdo test-retest cuando se mide a sujetos con alteraciones binoculares. Este parece ser el caso de los tests estudiados pues la repetibilidad empeoraba en bajos niveles de estereopsis (mayor grado de disparidad). Las variaciones en la repetibilidad en función de si se estudia en sujetos normales o en sujetos con deficiente estereoagudeza reduce intrínsecamente la validez de la prueba. En el presente estudio sólo se ha estudiado el comportamiento de los tests de estereopsis en una población con visión binocular normal, sin embargo, sería interesante estudiar la repetibilidad de los tests en una

población de sujetos con alteraciones binoculares para confirmar la tendencia que muestran los resultados de este estudio.

Leat *et al.*¹⁷⁰ estudiaron la repetibilidad del test de Randot y del test de Frisby para diferentes rangos de edad en sujetos con visión binocular normal. Para el grupo de sujetos de mayor edad (de 8 a 19 años, N=17), encontraron un percentil 95 de las diferencias entre medidas repetidas anormalmente bajo ($\pm 1,3$ seg arc), que es notoriamente menor al del presente estudio (Randot: ± 25 seg arc; Frisby: ± 35 seg arc). Esta discrepancia de resultados seguramente sea debida a que tal y como hacen ver los autores del estudio¹⁷⁰ prácticamente todos los sujetos de la escasa muestra alcanzaron sin dificultades el valor umbral para estos tests; mientras que en nuestro caso al contar con una muestra casi cuatro veces mayor se ha propiciado la aparición de una mayor gama de resultados.

CONCORDANCIA

La Tabla 28 muestra grandes diferencias entre los distintos tests estudiados. Excluyendo el test de Titmus con sus marcadas pistas monoculares, Marsh *et al.*¹⁵⁹ afirmaron que las diferencias entre los distintos tests podrían ser el resultado de factores tales como: (1) configuración y tamaño de las áreas de disparidad; (2) formas de las figuras, (3) distancia de observación, (4) gafas polarizadas vs. gafas rojo-verde, (5) tamaño de las figuras y (6) habilidad de los sujetos para identificar el área de disparidad.

Por otra parte, se ha comentado que la mayoría de los sujetos que no tienen anomalías oculares pueden discriminar diferencias de profundidad producidas por una disparidad relativa tan pequeña como 10 seg arc¹⁷¹. Sin embargo, la literatura que refleja esta evidencia no describe el tipo de test con el que se han conseguido esos niveles de estereoagudeza. Por ello, el rango y los saltos de disparidad en el test

deben considerarse factores importantes para comparar el comportamiento en los distintos tests. Un estereotest no sólo debería permitir medir un límite adecuado de estereoagudeza, sino que debería también permitir medir valores intermedios para poder discriminar la severidad de la alteración binocular.

Analizando los resultados de la Tabla 28, se puede decir que, tal y como ya mostró Hall¹⁷², el test TNO sería el test más difícil de pasar con un umbral de estereoagudeza medio próximo a 65 seg arc. El propio diseño del test TNO podría ser el responsable de esta mayor dificultad debido a distintos aspectos¹⁷²:

- Los saltos entre los mejores niveles de estereoagudeza son elevados (ver Tabla 25).
- Mientras que los demás tests de estereopsis estudiados sólo requieren la detección de una sensación de profundidad, el test TNO además necesita discriminación de la forma para poder identificar la dirección del sector circular que falta en el disco.
- Como ya se ha comentado, el TNO no presenta pistas monoculares.
- La fina textura del TNO hace que sea más difícil emparejar los elementos del estímulo vistos respectivamente por OD y OI.

En contra de lo esperado, el siguiente test en nivel de dificultad sería el de Titmus; sin embargo, quizá este resultado sea debido a que el límite medible con este test se sitúa en 40 seg arc y no en 20 seg arc como ocurre en las pruebas que se han mostrado como las más sencillas, el test de Randot y el de Frisby.

Rosner y Clift¹⁷³ compararon las medidas de estereoagudeza obtenidas con Frisby y TNO en 20 sujetos adultos de 23 a 37 años (media 27,4) con buena visión binocular y obtuvieron un coeficiente de correlación de Pearson $r=0,73$ ($p < 0,001$), tendiendo a dar

menores niveles de estereoagudeza TNO que Frisby. Hall¹⁷² midió los umbrales de estereoagudeza en 67 sujetos con visión binocular normal de 18 a 24 años utilizando distintos tests: Titmus, Frisby, TNO y prueba de las agujas. Encontraron que había una significativa, aunque pequeña correlación entre TNO y Frisby ($r=0,35$; $p<0,05$) y entre TNO y Titmus ($r=0,25$; $p<0,02$).

Es evidente que existen diferencias entre los distintos tipos de estereotests disponibles y entre los valores de umbral señalados para cada uno de ellos. Hay variación de opiniones sobre cuál es el estereotest más efectivo para su uso en la práctica clínica. Una característica deseable de los tests de estereoagudeza es que no contengan pistas monoculares, pues en esta situación la detección de la profundidad relativa sólo se hará en función de la disparidad entre las distintas partes de la imagen¹⁷². Si comparamos los tests **desde el punto de vista práctico**, el test de Frisby no necesita medios artificiales para evaluar la estereoagudeza, sin embargo, TNO, Randot y Titmus necesitan filtros anaglifos o polarizados para producir disparidad binocular, que podrían causar disociación parcial o introducir una rivalidad retiniana^{154, 165}. El tamaño angular de los estímulos es diferente para cada tests (y variable en el test de Frisby) lo que influye en el nivel de agudeza visual requerida para cada test. Por otra parte, para cambiar el grado de disparidad en el test de Frisby es necesario hacer variaciones poco prácticas de la distancia de observación, mientras que los demás tests se aplican a una distancia fija de 40 cm. Como ventaja, en el test de Frisby no se pueden memorizar o aprender las respuestas esperadas, lo que sí podría llegar a ocurrir en los otros tres tests evaluados, dado que sus elementos están fijos en una posición concreta. En definitiva, actualmente no es fácil decidir qué test es el más adecuado para la práctica clínica lo que muestra la necesidad de mejores diseños que reúnan en un solo test todas las cualidades deseables.

Conclusiones

La repetibilidad de las medidas de estereoagudeza fue baja. Con la presente configuración, los tests no parecen adecuados para detectar cambios pequeños en las capacidades visuales de un sujeto determinado ya que los posibles cambios reales que pudieran darse, quedarían ocultos dentro de los anchos intervalos asociados al error de medida. Además, la repetibilidad es peor cuanto menor sea el grado de estereopsis.

Así mismo, la reducida concordancia encontrada entre distintos tests indica que no pueden utilizarse de manera intercambiable. Esta falta de concordancia puede ser reflejo de los diferentes mecanismos subyacentes asociados a los distintos tipos de tests, al propio diseño del test o ser causada por la presencia de pistas no binoculares de profundidad. La peor concordancia aparece al comparar el TNO con los demás tests, donde además se ha visto que la concordancia disminuye a menor grado de estereopsis.

Conclusiones y perspectivas futuras

Conclusiones

En este trabajo se estudió la repetibilidad de las principales pruebas optométricas utilizadas en la evaluación clínica de la visión binocular y se analizó el grado de concordancia entre diferentes pruebas que evalúan la misma cualidad visual. En el apartado “1.2. *Aportaciones novedosas*” se reúnen en forma de tablas y se comentan brevemente los resultados clínicamente más útiles de este trabajo. Así mismo, al final de cada uno de los capítulos destinados a los distintos experimentos se han presentado las conclusiones particulares de cada parte del estudio dónde el lector podrá obtener más detalles sobre las conclusiones globales que aquí se exponen.

En este apartado se recogen las conclusiones más representativas que se pueden extraer de la presente investigación y que son las siguientes:

Respecto a la hipótesis general, se ha comprobado que:

- Efectivamente unas pruebas tienen distinto grado de repetibilidad que otras.
- Aunque el número de pruebas objetivas que se ha estudiado es bajo, sí se ha podido ver que el cover test para medir el grado de heteroforia muestra una repetibilidad mejor al resto de los métodos de su grupo. Respecto al grupo de pruebas para medir la respuesta acomodativa donde se acumulaban el resto de las pruebas objetivas estudiadas, en concreto tres de los cuatro métodos eran objetivos, sólo una de las pruebas objetivas, la retinoscopía Nott, ha mostrado una repetibilidad mejor a los restantes métodos. Son necesarios más estudios para

poder confirmar que la repetibilidad de las pruebas objetivas es mayor que el de las subjetivas.

- En cada grupo de pruebas se ha comprobado que el grado de concordancia entre parejas de métodos varía de unas parejas a otras, confirmando que tiende a ser mayor cuanto mejor es la repetibilidad de los dos métodos comparados.

A continuación se incluyen algunas conclusiones globales para cada grupo de pruebas que miden la misma variable:

- **Medida de la amplitud de acomodación:**

La repetibilidad del método de las lentes negativas fue mucho mejor a la presentada por los métodos de acercamiento y alejamiento, siendo la anchura de su intervalo de acuerdo al 95% prácticamente la mitad del presentado por los otros dos métodos.

La concordancia entre las tres pruebas fue baja, en concreto, la medida de la amplitud de acomodación con la técnica de acercamiento como promedio fue casi 5 D superior a la obtenida con lentes negativas y aproximadamente 2 D mayor que la medida con el método de acercamiento, confirmándose la tendencia encontrada en investigaciones previas.

- **Medida de la respuesta acomodativa:**

Los resultados han confirmado que la retinoscopia Nott es la técnica más adecuada para medir la respuesta acomodativa.

A la hora de prescribir una adición positiva como tratamiento de una anomalía acomodativa o de vergencias se deberá tener en cuenta que la concordancia

entre las distintas técnicas estudiadas es baja, con la excepción de las dos técnicas de retinoscopia dinámica estudiadas, donde la discrepancia fue bastante aceptable ($CDC = \pm 0,64 D$).

- **Medida de la desviación ocular:**

El cover test se ha mostrado como una técnica muy repetible en contraposición de lo que ocurre con la técnica de von Graefe, donde los cambios achacables a un error de medida pueden superar las 6Δ . Este estudio ha confirmado que el test de Thorington modificado es la prueba subjetiva de medida de la desviación más repetible y junto con el cover test (punto de neutralización) son las pruebas recomendadas para la práctica clínica.

Todas las técnicas estudiadas presentaron mejor repetibilidad y concordancia al aplicarlas en visión lejana que en visión próxima.

Tal y como era de esperar debido a las notorias diferencias metodológicas, la concordancia entre los distintos métodos de medida de la desviación fue baja, especialmente en visión próxima, por lo que se recomienda utilizar la misma prueba para las sucesivas revisiones de un determinado paciente. La mejor concordancia a esta distancia^{***} la presentaron la técnica de von Graefe y la de Thorington modificado, con la menor amplitud de su intervalo de acuerdo al 95%.

Tanto en visión lejana como en cerca el grado de repetibilidad no variaba en función de la magnitud de la desviación. Sin embargo, la concordancia entre métodos objetivos y subjetivos de medida de la foria horizontal sí empeoraba al aumentar el valor de la desviación. Se necesitan más estudios para poder caracterizar mejor este comportamiento.

^{***} Se excluye de la discusión el buen acuerdo que aparece entre las dos modalidades de cover test.

- **Medida de las reservas fusionales:**

La medida del PPC es una prueba con buena repetibilidad.

Sin excepción, la repetibilidad ha sido mejor en la medida de la divergencia que en la de la convergencia.

En términos generales, la diferencia media en base externa (convergencia) cuadruplicaba a la de base interna (divergencia) (2Δ vs. $0,5\Delta$) y el coeficiente de repetibilidad en base externa duplicaba a su correspondiente de base interna.

Los valores medidos de emborronamiento y rotura siempre fueron mayores utilizando diasporámetros, pero se recuperó más rápidamente la fusión (mayor recobro) utilizando la barra de prismas.

No hubo una buena correlación entre el resultado de la medida de la convergencia con barra de prismas o diasporámetro respecto a la obtenida con la prueba del PPC.

- **Medida de la estereoagudeza**

La repetibilidad de los test de estereopsis estudiados fue baja en el rango de valores de disparidad de fijación por encima de 50 seg arc, por lo que no parecen adecuados para detectar pequeños cambios en el nivel de estereoagudeza. Son necesarios más estudios para poder caracterizar mejor este comportamiento.

Como se esperaba dada las notorias diferencias de diseño, la concordancia entre los distintos tests fue baja, especialmente al comparar el TNO con los demás estereotests, siendo la concordancia peor cuanto menor era el grado de estereoagudeza (mayor disparidad).

- **Otras pruebas:**

La medida de la acomodación relativa positiva (ARP) mostró ser muy poco repetible, posiblemente debido al alto grado de subjetividad inherente a la prueba. También parecen pruebas poco repetibles la medida de la flexibilidad acomodativa, tanto mono como binocular, y la de flexibilidad de vergencias. Son pruebas que han reflejado tener un error de medida asociado tan elevado como para no permitir detectar cambios reales en la habilidad medida por lo que su utilidad clínica está limitada.

Por último, se considera interesante incluir en este apartado algunas conclusiones generales sobre el comportamiento de las pruebas clínicas estudiadas en relación con tres aspectos interesantes: índices estadísticos que se han mostrado más útiles, efecto aprendizaje y enmascaramiento.

- **Índices estadísticos**

Respecto a la experiencia adquirida al aplicar las técnicas de análisis estadístico a este estudio, se ha visto que al estudiar la repetibilidad de este tipo de pruebas, desde el punto de vista de la práctica clínica a la hora de extraer conclusiones se debe tener en cuenta sobre todo el intervalo de acuerdo al 95% (CDR) pues las diferencias medias (DM) encontradas, incluso teniendo en cuenta sus correspondientes intervalos de confianza^{†††}, no alcanzaron valores que pudieran considerarse clínicamente relevantes.

En cuanto al estudio de la concordancia entre métodos, de nuevo, en general, el parámetro más importante a la hora de establecer el grado de acuerdo entre

^{†††} En las tablas de resultados no se han incluido los valores de los intervalos de confianza encontrados para la diferencia media y para los límites del intervalo de acuerdo al 95% debido a que no añadían información adicional a las conclusiones extraíbles del estudio.

distintas pruebas volvió a ser el intervalo de acuerdo al 95% (CDC). Sin embargo, en algunos grupos de pruebas la diferencia media (DM) sí aportó datos interesantes, siendo este el caso de la medida de la amplitud de acomodación, de la desviación ocular y de la estereoagudeza.

- **Efecto aprendizaje**

En la medida de la flexibilidad, tanto acomodativa como de vergencias, se manifestó un efecto aprendizaje al repetir la prueba, aumentando el número de ciclos obtenidos en sucesivas sesiones de medida. Ello hace que aquellos clínicos que deseen utilizar la evaluación de la flexibilidad dentro de una batería de pruebas de *screening* visual, deberán tener en cuenta que es posible que se produzcan muchos falsos positivos si sólo se aplica la prueba en una ocasión.

En el estudio de repetibilidad, las diferencias medias (DM) encontradas no siguieron un comportamiento único, por ejemplo, las medidas de amplitud de acomodación y retraso acomodativo reflejaron que como promedio los resultados de la segunda tanda fueron menores que en la primera; sin embargo, en la medida de la flexibilidad y de los rangos de vergencia^{†††} hubo tendencia a obtener mayores resultados en la segunda sesión. Por ello se puede concluir que las DMs entre primera y segunda sesión de medidas no pueden atribuirse exclusivamente a un efecto aprendizaje, situación en la que la media de la segunda sesión superaría siempre a la primera.

^{†††} Todas las medidas de emborronamiento (E), rotura (Ro) y recobro (Re) reflejaron como promedio mayor amplitud en la segunda sesión de medidas que en la primera excepto en: E VFP en lejos con diaporámetro y Ro / Re VFN en cerca tanto con barra de prismas como con diaporámetro.

- **Enmascaramiento del examinador**

Los resultados de repetibilidad intra-examinador encontrados, especialmente para las pruebas objetivas, podría estar algo sesgados por la incapacidad de enmascarar totalmente al observador que podría recordar esporádicamente los resultados de algún sujeto. Sin embargo, el alto número de sujetos examinados, combinado con el periodo entre sesiones de medidas (media: 7,01; DE=1,60 días) y, por supuesto, la ocultación de los datos de la primera tanda de medidas al llevar a cabo la segunda tanda, hacen que el hecho de que el examinador pueda recordar los resultados de medidas previas sea extremadamente difícil. Por ello, creemos que las deficiencias en el enmascaramiento no han afectado a los resultados del estudio de forma apreciable.

Propuestas para investigaciones futuras

Fruto de la revisión minuciosa de los estudios en este campo y de las reflexiones sobre los resultados obtenidos surgen las siguientes propuestas para continuar nuestra investigación en esta línea de trabajo:

- En esta tesis doctoral se ha hecho una revisión y estudio de la repetibilidad intra-examinador y podría continuarse con la investigación de la **repetibilidad inter-examinador**. Cabe esperar que cuando se estudie la repetibilidad inter-examinador el error sea mayor que en la repetibilidad intra-examinador en pruebas objetivas y casi igual en subjetivas. Porque, según se comenta en la introducción, si las medidas de una determinada prueba son tomadas por varios examinadores se estaría introduciendo otra fuente de variabilidad y cabe esperar un aumento del error de la medida. Posiblemente este aumento del error no será importante si la

intervención del examinador en la medida es pequeña (pruebas subjetivas) y será más elevado en las pruebas objetivas.

- Las conclusiones de esta investigación se basan en el análisis de los resultados extraídos de una muestra de sujetos adultos jóvenes (18 a 32 años; media: 19,74; DE: 2,5 años). Dado que cabe esperar que algunas de las conclusiones podrían variar si se estudiaran otros rangos de edad, se pretende continuar el estudio con **poblaciones dentro de otros rangos de edad** (niños, présbitas, ancianos). Las pruebas de visión binocular estudiadas en este trabajo son especialmente útiles al evaluar a sujetos con síntomas astenópicos asociados generalmente a altas demandas visuales en visión próxima como las que soportan la población de la cual se extrajo la muestra de este estudio. Se pretende ampliar el estudio a poblaciones más generales en cuanto a rango de edad y demandas visuales.
- Hay que recordar que el estudio se ha realizado en condiciones más o menos ideales, con profesionales motivados y entrenados en la aplicación de las pruebas, y las conclusiones halladas puede que no fueran las mismas si el examinador contara con menor experiencia o fuera estudiante. Este factor puede tener especial relevancia en las pruebas de tipo objetivo donde la pericia del examinador es la que determina el resultado de la medida. Por ejemplo, podría ser interesante estudiar la repetibilidad del cover test o la retinoscopía MEM o Nott cuando las **pruebas fueran aplicadas por un examinador no experimentado**.
- Es deseable que para obtener un mayor alcance de los resultados extraídos de un estudio, los sujetos incluidos en el mismo sean una muestra representativa de la población a la que se piensa aplicar la medida y cubrir un amplio rango de valores de la variable que se desea medir, para poder evaluar la fiabilidad durante todo su recorrido. Por otra parte, quedan varios aspectos de la visión binocular que no se

han evaluado con las pruebas estudiadas en la presente investigación como la disparidad de fijación, las ciclodesviaciones, la correspondencia sensorial, la fijación, la motilidad ocular, etc. El estudio de muchas de ellas sólo es interesante cuando se lleva a cabo en poblaciones especiales como estrábicos, amblíopes, sujetos con altas heteroforias o niños con problemas de lectura. Una futura investigación se destinará a estudiar la repetibilidad y el grado de concordancia de algunas de las pruebas incluidas en este estudio y de otras pruebas que sirvan para evaluar los aspectos citados anteriormente cuando se **aplican en poblaciones con alteraciones específicas**.

- Por último, también dentro de nuestra línea de trabajo, en futuras investigaciones, nos proponemos determinar el **valor diagnóstico de las distintas pruebas** utilizadas en evaluación de anomalías acomodativas y de las vergencias, bien aisladamente o bien en grupos de pruebas que pretendan diagnosticar cada anomalía. Los índices que permiten definir el valor diagnóstico de las pruebas son fundamentalmente: sensibilidad, especificidad y valor predictivo. Cuando se conocen los valores de estos índices se puede llevar a cabo con mayor certidumbre la difícil tarea de diagnosticar la anomalía que presente el paciente. Un paciente puede mostrar una serie de síntomas y signos poco específicos, que podrían corresponder a diferentes anomalías, lo que provoca una situación de incertidumbre que debe ser disminuida con la utilización de pruebas validas y fiables que permitan hacer correctamente el diagnóstico diferencial.

Referencias

1. Last JM. Diccionario de epidemiología. Barcelona: Salvat; 1989.
2. Latour J, Abaira V, Cabello JB, López J. Las mediciones clínicas en cardiología: validez y errores de medición. *Rev Esp Cardiol* 1997;50:117-128.
3. Argimon JM, Jiménez J. Métodos de investigación clínica y epidemiológica. 3 ed. Madrid: Elsevier; 2004.
4. Bullimore MA. The way you do the things you do. *Optom Vis Sci* 1998;75:707.
5. Reeves BC, Hill AR, Aspinall PA. The clinical significance of change. *Ophthalm Physiol Opt* 1987;7:441-446.
6. Pellizzer S, Siderov J. Assessment of vergence facility in a sample of older adults with presbyopia. *Optom Vis Sci* 1998;75:817-821.
7. Rouse MW, Borsting E, Deland PN. Reliability of binocular vision measurements used in the classification of convergence insufficiency. *Optom Vis Sci* 2002;79:254-264.
8. McKenzie KM, Kerr SR, Rouse MW, Deland PN. Study of accommodative facility testing reliability. *Am J Optom Physiol Optics* 1987;64:186-194.
9. Rouse MW, Deland PN, Chous R, Determan TF. Monocular accommodative facility testing reliability. *Optom Vis Sci* 1989;66:72-77.
10. Rouse MW, DeLand PN, Mozayani S, Smith JP. Binocular accommodative facility testing reliability. *Optom Vis Sci* 1992;69:314-319.

11. Elliott DB. Clinical procedures in primary eye care. Edinburgh: Butterworth Heinemann; 2003.
12. Eskridge JB, Amos JF, Bartlett JD. Clinical procedures in optometry. Philadelphia: Lippincott Company; 1991.
13. Zadnik K. The ocular examination. Measurements and findings. Philadelphia: WB Saunders Company; 1997.
14. Scheiman M, Wick B. Clinical management of binocular vision. Heterophoric accommodative and eye movement disorders. Philadelphia: Lippincott Williams & Wilkins; 2002.
15. Carlson NB, Kurtz D. Clinical procedures for ocular examination. 3 ed. New York: Mc Graw Hill; 2004.
16. Benjamin WJ. Borish's clinical refraction. 2 ed. St. Louis: Butterworth Heinemann Elsevier; 2006.
17. Lloyd M. Lies, statistics and clinical significance. J Br Contact Lens Ass 1992;15:67-70.
18. Reeves BC, Wood JM, Hill AR. Vistech VCTS 6500 Charts-within and between-session reliability. Optom Vis Sci 1991;68:728-737.
19. Antona B, Barra F, González E, Sánchez I, Barrio A. Fiabilidad y validez de pruebas optométricas. Libro actas de la I Jornada Complutense de Investigación en Optometría y Ciencias de la Visión. Madrid: Universidad Complutense de Madrid; 2005:107-109.
20. Muñiz J. Teoría clásica de los tests. Madrid: Pirámide; 1992.

21. Hernández I, Porta M, Miralles M, García F, Bolumar F. La cuantificación de la variabilidad en las observaciones clínicas. *Med Clin (Barc)* 1990;95:424-429.
22. Hulley SB, Cummings SR. Planificación de las mediciones: precisión y exactitud. En: Hulley SB, Cummings SR (eds), *Diseño de la investigación clínica Un enfoque epidemiológico*. Barcelona: Doyma; 1993:35-46.
23. Elliott DB. Supplementary clinical test of visual function. En: Zadnik K (ed), *The ocular examination: measurements and findings*. Philadelphia: Saunders; 1997.
24. Howarth PA, Heron G. Repeated measures of horizontal heterophoria. *Optom Vis Sci* 2000;77:616-619.
25. Hulley SB, Cummings SR. *Designing clinical research: An epidemiologic approach*. Baltimore: Williams & Wilkins; 1988.
26. Cepeda MS, Pérez A. Estudios de concordancia. Intercambiabilidad en sistemas de medición. En: Ruíz A, Morillo LE (eds), *Epidemiología clínica Investigación clínica aplicada*. Bogotá: Editorial Médica Panamericana; 2004:293-307.
27. McClelland JK, Saunders KJ. The repeatability and validity of dynamic retinoscopy in assessing the accommodative response. *Ophthal Physiol Opt* 2003;23:243-250.
28. Bland J, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1-8476 1986;307-310.
29. Riegelman RK. *Studying a study and testing a test*. Philadelphia: Lippincott Williams & Wilkins; 2000:127-176.

30. Levine S, Ciuffreda KJ, Selenow A. Clinical assessment of accommodative facility in symptomatic and asymptomatic individuals. *J Am Optom Assoc* 1985;56:286-290.
31. Zadnik K, Mutti DO, Adams AJ. The repeatability of measurement of the ocular components. *Invest Ophthal Vis Sci* 1992;33:2325-2333.
32. D'Augostino RB, Stevens MA. Tests for the normal distribution. New York: Marcel Dekker; 1986:367-419.
33. Rosenfield M, Gilmartin B. Effect of target proximity on the open-loop accommodative response. *Optom Vis Sci* 1990;67:74-79.
34. Hokoda SC. General binocular dysfunction in an urban optometric clinic. *J Am Optom Assoc* 1985;56:560-562.
35. Rosenfield M. Accommodation. En: Zadnik K (ed), *The ocular examination; measurements and findings*. Philadelphia: WB Saunders; 1997.
36. London R. Accommodation. En: Barresi BJ (ed), *Ocular assessment The manual of diagnosis for office practice*. Boston: Butterworth-Heinemann; 1984:123-130.
37. Wick B, Hall P. Relation among accommodative facility, lag, and amplitude in elementary school children. *Am J Optom Physiol Opt* 1987;64:593-598.
38. Goss D. Clinical accommodation testing. *Current Opinion in Ophthalmology* 1992;3:78-82.
39. Woehrle MB, Peters RJ, Frantz KA. Accommodative amplitude determination: Can we substitute the pull-away for the push-up method? *J Optom Vis Develop* 1997;28:246-249.

40. Brozek J, Simonson E, Bushard JW, Peterson HJ. Effects of practice and the consistency of repeated measurements of accommodation and vergence. *Am J Ophthalmol* 1948;31:191-198.
41. Rosenfield M, Cohen AS. Repeatability of clinical measurements of the amplitude of accommodation. *Ophthal Physiol Opt* 1996;16:247-249.
42. Chen AH, O'Leary DJ. Validity and repeatability of the modified push-up method for measuring the amplitude of accommodation. *Clin Exp Optom* 1998;81:63-71.
43. Atchison DA, Capper EJ, McCabe KL. Critical subjective measurement of amplitude of accommodation. *Optom Vis Sci* 1994;71:699-706.
44. Fitch R. Procedural effects on the manifest human amplitude of accommodation. *Am J Optom Arch Am Acad Optom* 1971;48:918-926.
45. Ostrin LA, Glasser A. Accommodation measurements in a prepresbyopic and presbyopic population. *J Cataract Refract Surg* 2004;30:1435-1444.
46. Rambo VC, Sangal SP. A study of the accommodation of the people of India. With further notes on the development of presbyopia at different ages in different peoples. *Am J Ophthalmol* 1960;49:993-1004.
47. Kragha IKOK. Measurement of amplitude of accommodation. *Ophthal Physiol Opt* 1989;9:342-343.
48. Wold JE, Hu A, Chen S, Glasser A. Subjective and objective measurement of human accommodative amplitude. *J Cataract Refract Surg* 2003;29:1878-1888.

49. Win-Hall DM, Ostrin LA, Kasthurirangan S, Glasser A. Objective accommodation measurement with the Grand Seiko and Haidinger Coincidence Refractometer. *Optom Vis Sci* 2007;84:879-887.
50. Hokoda SC, Ciuffreda KJ. Measurement of accommodative amplitude in amblyopia. *Ophthalm Physiol Opt* 1982;15:231-232.
51. Rosenfield M, Cohen AS. Push-up amplitude of accommodation and target size. *Ophthalmic Physiol Opt* 1995;15:231-232.
52. Jackson TW, Goss DA. Variation and correlation of clinical tests of accommodative function in a sample of school-age children. *J Am Optom Assoc* 1991;62:857-866.
53. Wick B, Yothers TL, Bai-Chuan J, Morse SE. Clinical testing of accommodative facility: Part I. A appraisal of the literature. *Optometry* 2002;73:11-23.
54. Jiménez R, González R, Jiménez JR. Actualización optométrica pediátrica: función acomodativa (II). *Gaceta Optica* 2007;410:10-15.
55. Hennessey D, Iosue RA, Rouse MW. Relation of symptoms to accommodative infacility of school-aged children. *Am J Optom Physiol Opt* 1984;61:177-183.
56. Siderov J, DiGuglielmo L. Binocular accommodative facility in prepresbyopic adults and its relation to symptoms. *Optom Vis Sci* 1991;68:49-53.
57. Wick B, Gall R, Yothers T. Clinical testing of accommodative facility: Part III. Masked assessment of the relation between visual symptoms and binocular test results in school children and adults. *Optometry* 2002;73:173-180.

58. Saladin JJ. Phorometry and Stereopsis. En: Benjamin WJ (ed), Borish's clinical refraction. St. Louis: Butterworth Heinemann Elsevier; 2006:899-960.
59. Kedzia B, Pieczyrak D, Tondel G, Maples WC. Factors affecting the clinical testing of accommodative facility. *Ophthal Physiol Opt* 1999;19:12-21.
60. Radhakrishnan H, Allen PM, Charman WN. Dynamics of accommodative facility in myopes. *Invest Ophthal Vis Sci* 2007;48:4375-4382.
61. Siderov J, Johnston AW. The importance of test parameters in the clinical assessment of accommodative facility. *Optom Vis Sci* 1990;67:551-557.
62. Zellers JA, Alpert TL, Rouse MW. A review of the literature and a normative study of accommodative facility. *J Am Optom Assoc* 1984;1984.
63. Scheiman M, Herzberg H, Frantz K, Margolies M. Normative study of accommodative facility in elementary school children. *Am J Optom Physiol Opt* 1988;65:127-133.
64. Evans B. *Visión binocular*. Barcelona: Masson; 2006.
65. Schor C. The influence of interactions between accommodation and convergence on the lag of accommodation. *Ophthal Physiol Opt* 1999;19:134-150.
66. Goss DA, Groppe P, Domínguez L. Comparison of MEM retinoscopy & Nott retinoscopy & their interexaminer repeatabilities. *J Behav Optom* 2005;16:149-155.
67. Cacho MP, García-Muñoz A, García-Bernabeu JR, López A. Comparison between MEM and Nott dynamic retinoscopy. *Optom Vis Sci* 1999;76:650-655.
68. Rosenfield M, Ciuffreda KJ. Accommodative responses to conflicting stimuli. *J Opt Soc Am A* 1991;8:422-427.

69. Adams CW, Johnson CA. Steady-state and dynamic response properties of the Mandelbaum effect. *Vis Res* 1991;31:751-760.
70. Rosenfield M, Portello JK, Blustein GH, Jang C. Comparison of clinical techniques to assess the near accommodative response. *Optom Vis Sci* 1996;73:382-388.
71. Goodson RA, Afanador AJ. The accommodative response to the near point crossed cylinder test. *Optom Wkly* 1974;65:1138-1140.
72. Locke LC, Somers W. A comparison study of dynamic retinoscopy techniques. *Optom Vis Sci* 1989;66:540-544.
73. McKee GW. Reliability of monocular estimate method retinoscopy. *Optom Mon* 1981;72:30-31.
74. Chat SW, Edwards MH. Clinical evaluation of the Shin-Nippon SRW-5000 autorefractor in children. *Ophthal Physiol Opt* 2001;21:87-100.
75. Mallen EAH, Wolffsohn JS, Gilmartin B, Tsujimura S. Clinical evaluation of the Shin-Nippon SRW-5000 autorrefractor in adults. *Ophthal Physiol Opt* 2001;21:101-107.
76. Wolffsohn JS, Gilmartin B, Mallen EA, Tsujimura S. Continuous recording of accommodation and pupil size using the Shin-Nippon SRW-5000 autorrefractor. *Ophthal Physiol Opt* 2001;21:108-113.
77. Nakasuka C, Hasebe S, Nonaka F, Ohtsuki H. Accommodative lag under habitual seeing conditions: comparison between adult myopes and emmetropes. *Jpn J Ophthalmol* 2003;47:291-298.

-
78. Wolffsohn JS, Gilmartin B, Mallen EA, Tsujimura S. Simultaneous continuous recording of accommodation and pupil size using the modified Shin-Nippon SRW-5000 autorrefractor. *Ophthal Physiol Opt* 2004;24:142-147.
79. Nakasuka C, Hasebe S, Nonaka F, Ohtsuki H. Accommodative lag under habitual seeing conditions: comparison between myopic and emmetropic children. *Jpn J Ophthalmol* 2005;49:189-194.
80. Whitefoot H, Charman WN. Dynamic retinoscopy and accommodation. *Ophthal Physiol Opt* 1992;12.
81. Campbell CE, Benjamin WJ, Howland HC. Objective refraction: retinoscopy, autorefractometry and photorefractometry. En: Benjamin WJ (ed), *Borish's clinical refraction*. Philadelphia: Saunders; 1998:559-628.
82. Kimura S, Hasebe S, Ohtsuki H. Systematic measurement errors involved in over-refraction using an autorrefractor (Grand-Seiko WV-500): is measurement of accommodative lag through spectacle lenses valid? *Ophthal Physiol Opt* 2007;27:281-286.
83. García A, Cacho MP. MEM and Nott dynamic retinoscopy in patient with disorders of vergence and accommodation. *Ophthal Physiol Opt* 2002;22:214-220.
84. Goss DA. Ocular accommodation, convergence and fixation disparity: A manual of clinical analysis. 2 ed. Boston: Butterworths-Heinemann; 1995:11-12, 70.
85. Goss D, Zhai H. Clinical and laboratory investigations of the relationship of accommodation and convergence function with refractive error. *Documenta Ophthalmologica* 1994;86:349-380.

86. Cline D, Hofstetter HW, Griffin JR. Dictionary of visual science. 4 ed. Boston: Butterworth-Heinemann; 1989:313-314.
87. Griffin JR. Binocular anomalies: Procedures for vision therapy. 2 ed. Chicago: Professional Press; 1982:140-147.
88. Fisher SK, Ciuffreda KJ, Tannen B. Stability of tonic vergence. Invest Ophthal Vis Sci 1988;29:1577-1581.
89. Bernal J, Rosenfield M. Effect of heterophoria measurement technique on the clinical accommodative convergence to accommodation. Optometry 2006;77:229-234.
90. Wong EP, Fricke TR, Dinardo C. Interexaminer repeatability of a new, modified Prentice card compared with established phoria tests. Optom Vis Sci 2002;79:370-375.
91. Johns HA, Manny RE, Fern K, Hu Y. The intraexaminer and interexaminer repeatability of the alternate cover test using different prism neutralization endpoints. Optom Vis Sci 2004;81:939-946.
92. Schroeder TL, Rainey BB, Goss DA, Grosvenor TP. Reliability of and comparison among methods of measuring dissociated phoria. Optom Vis Sci 1996;73:389-397.
93. Rainey BB, Schroeder TL, Goss DA, Grosvenor TP. Inter-examiner repeatability of heterophoria test. Optom Vis Sci 1998;75:719-726.
94. Schor CM, Ciuffreda KJ. Vergence Eye Movements. Basic and clinical aspects. Boston: Butterworth; 1983.
95. Scobee RG, Green EL. Tests for heterophoria-reliability of tests, comparisons between tests, and effect of changing teststing conditions. Trans Am Acad Ophthalmol Otolaryngol 1947;51:179-197.

96. Hirsch MJ, Bing LB. The effect of testing method on values obtained for phoria at forty centimeters. *Am J Optom Arch Am Acad Optom* 1948;25:407-416.
97. Doherty BA, Geraci MJ. Comparative phoric measurements. *Optom Wkly* 1954;45:1479-1981.
98. Morgan MW. The reliability of clinical measurements with special reference to distance heterophoria. *Am J Optom Arch Am Acad Optom* 1955;32:167-179.
99. Morris FM. The influence of kinesthesia upon near heterophoria measurements. *Am J Optom Arch Am Acad Optom* 1960;37:327-351.
100. Lam AKC, Lam A, Charm J, Wong K. Comparison of near heterophoria tests under varying conditions on an adult sample. *Ophthal Physiol Opt* 2005;25:162-167.
101. Casillas E, Rosenfield M. Comparison of subjective heterophoria testing with a phoropter and trial frame. *Optom Vis Sci* 2006;83.
102. Hirsch MJ. Clinical investigation of a method of testing phoria at forty centimeters. *Am J Optom Arch Am Acad Optom* 1948;25:492-495.
103. Weymouth FW. An experimental comparison of three common methods of measuring heterophoria. *Am J Optom Arch Am Acad Optom* 1963;40:497-503.
104. Soderberg DC. An evaluation in the use of Maddox rod. *J Am Optom Assoc* 1968;39:472-478.
105. Daum KM. Analysis of seven methods of measuring the angle of deviation. *Am J Optom Physiol Opt* 1983;60:46-51.
106. Calvin H, Rupnow P, Grosvenor T. How good is the estimated cover test at predicting the von graefe phoria measurement? *Optom Vis Sci* 1996;73:701-706.

107. Thompson JT, Guyton DL. Ophthalmic prisms. Measurement errors and how to minimize them. *Ophthalmology* 1983;90:204-210.
108. Charman WN, Heron G. Fluctuations in accommodation: a review. *Ophthalmic Physiol Opt* 1988;8:153-164.
109. Ludvigh E. Amount of eye movement objectively perceptible to the unaided eye. *Am J Ophthalmol* 1949;32:649-650.
110. Fogt N, Baughman BJ, Good G. The effect of experience on the detection of small eye movements. *Optom Vis Sci* 2000;77:670-674.
111. Burian HM. Burian-von Noorden's binocular vision and ocular motility: Theory and management of strabismus. St Louis: Mosby; 1980:182-188.
112. Sparks BI. Phoria variation secondary to cover test technique at near. *Optometry* 2002;73:51-54.
113. Clack TE, Reynolds W, Coffey B. The effects of varying examiner position on alternating cover test results. *Optometry* 2003;74:233-239.
114. Owens DA, Leibowitz HW. The fixation point as a stimulus for accommodation. *Vis Res* 1975;15:1161-1163.
115. Benjamin W. Borish' clinical refraction. Philadelphia: WS Saunders; 1998:1142-1159.
116. Frantz KA, Scharre JE. Comparison of diasporameter fixation disparity curves as measured with and without the phorofter. *Optom Vis Sci* 1990;67:117-122.
117. North RV, Henson DB, Smith TJ. Influence of proximal, accommodative and disparity stimuli upon the vergence system. *Ophthal Physiol Opt* 1993;13:239-243.

118. Rosenfield M. Tonic vergence and vergence adaptation. *Optom Vis Sci* 1997;74:303-328.
119. Daum KM. Vergence amplitude. En: Eskridge JB, Amos JF, Bartlett JD (eds), *Clinical procedures in optometry*. Philadelphia: Lippincott; 1991:91-98.
120. Sheedy JE, Saladin JJ. Phoria, vergence and fixation disparity in oculomotor problems. *Am J Optom Physiol Opt* 1977;54:474-478.
121. Rosenfield M, Ciuffreda KJ, Ong E. Vergence adaptation and the order of clinical vergence range testing. *Optom Vis Sci* 1995;72:219-223.
122. Goss DA. Effect of test sequence on fusional vergence ranges. *N Engl J Optom* 1995;47:39-42.
123. Feldman JM, Cooper J, Carniglia P, Schiff FM, Skeete JN. Comparison of fusional ranges measured by Risley prisms, vectograms and computer orthopter. *Optom Vis Sci* 1989;66:375-382.
124. Penisten DK, Hofstetter HW, Goss DA. Reliability of rotary prism fusional vergence ranges. *Optometry* 2001;72:117-122.
125. Ciuffreda MA, Ciuffreda KJ, Wang B. Repeatability and variability of near vergence ranges. *J Behav Optom* 2006;17:39-46.
126. Scheiman M, Mitchell GL, Cotter S, et al. A randomized clinical trial of treatment for convergence insufficiency in children. *Arch Ophthalmol* 2005;123:14-24.
127. Adler PM, Cregg M, Viollier AJ, Woodhouse M. Influence of target type and RAF rule on the measurement of near point of convergence. *Ophthal Physiol Opt* 2007;27:22-30.

128. Cardona JC, Pérez MA, García JA. Analisis comparativo de distintos métodos de medida del punto próximo de convergencia. *Archivos Optométricos* 1999;3:90-99.
129. Ciuffreda KJ. Near point of convergence as a function of target accommodative demand. *Opt J Rev Optom* 1974;111:9-10.
130. Siderov J, Chiu SC, Waugh SJ. Differences in the near point of convergence with target type. *Ophthal Physiol Opt* 2001;21.
131. Adler P. Optometric evaluation of binocular vision anomalies. En: Doshi S, Evans BJW (eds), *Binocular vision and orthoptics*. Oxford: Butterworth-Heinemann; 2001:101-109.
132. Sethi B. Vergence adaptation: a review. *Doc Ophthalmol* 1986;63:247-263.
133. Stephens GL, Jones R. Horizontal fusional amplitudes after adaptation to prism. *Ophthal Physiol Opt* 1990;10:25-28.
134. Firth AY. Vergence adaptation: a phenomenon of normal binocular vision. *Br Ir Orthopt J* 2005;2:3-7.
135. Neely JC. The RAF near point rule. *Br J Ophthal* 1956;40:636-637.
136. Morgan M. Analysis of clinical data. *Am J Optom Arch Am Acad Optom* 1944;21:477-491.
137. Wesson MD. Normalization of prims bar vergences. *Am J Optom Physiol Opt* 1982;59:628-633.
138. Sheedy JE, Saladin JJ. Validity of diagnostic criteria and case analysis in binocular vision disorders. En: Schor CM, Ciuffreda KJ (eds), *Vergence movements: basic and clinical aspects*. Boston: Butterworth; 1983:517-540.

139. Tannen B, Ciuffreda KJ. Eye movement basic for the clinician. St. Louis: Mosby Year-Book; 1995.
140. Scheiman M, Gallaway M, Frantz KA, et al. Nearpoint of convergence: test procedure, target selection, and normative data. *Optom Vis Sci* 2003;80:214-225.
141. Kumar AN, Han Y, Garbutt S, Leigh RJ. Properties of anticipatory vergence responses. *Invest Ophthal Vis Sci* 2002;43:2626-2632.
142. Scheiman M, Herzberg H, Frantz K. A normative study of step vergence in elementary schoolchildren. *J Am Optom Assoc* 1989;60:276-280.
143. Burian HM. Fusional eye movements: role of peripheral retinal stimuli. *Arch Ophthalmol* 1939;21:486-491.
144. Tanimoto N, Y Takagi M, Bando T, et al. Central and peripheral visual interactions in disparity-induced vergence eye movements. I. Spatial interaction. *Invest Ophthal Vis Sci* 2004;45:1132-1138.
145. Gall R, Wick B, Bedell H. Vergence facility: establishing clinical utility. *Optom Vis Sci* 1998;75:731-742.
146. Gall R, Wick B, Bedell H. Vergence facility and target type. *Optom Vis Sci* 1998;75:727-730.
147. Hoffman LC, Rouse M. Referral recommendations for binocular function and/or developmental perceptual deficiencies. *J Am Optom Assoc* 1980;51:119-126.
148. Waugh SJ. The value of measuring convergence accommodation and rapid prism adaptation clinically. Unpublished Master`s thesis. University of Melbourne; 1986.

149. Sheedy JE, Saladin JJ. Exophoria at near in presbyopia. *Am J Optom Physiol Opt* 1975;52:474-481.
150. Rosenfield M, Ciuffreda KJ. Accommodative convergence and age. *Ophthal Physiol Opt* 1990;10:403-404.
151. Fielder AR, Moseley MJ. Does stereopsis matter in humans? *Eye* 1996;10:233-238.
152. Jiménez R, Jiménez JR, González R, Pozo AM. Importancia de la estereopsis en la vida cotidiana. *Gaceta Optica* 2006;404:10-12.
153. Broadbent H, Westall C. An evaluation of techniques for measuring stereopsis in infants and young children. *Ophthal Physiol Opt* 1990;10:3-7.
154. Simons K. A comparison fo the Frisby, Random-Dot E, TNO, and Randot Circles stereotests in screening and office use. *Arch Ophthalmol* 1981;99:446-452.
155. Morris H, O'Connor AR, Stephenson MG, Mitxhell M, Price G, Anderson S. Clinical assessment of stereopsis and its functional significance. *Br Ir Orthopt J* 2005;2:8-14.
156. Cooper J, Warshowsky J. Lateral displacement as a response cue in the Titmus Stereo Test. *Am J Optom Physiol Optics* 1977;54:537-541.
157. Julesz B. Foundations of cyclopean perception. Chicago: University Chicago Press; 1971.
158. Fricke TR, Siderov J. Stereopsis, stereotests, and their relation to vision screening and clinical practice. *Clin Exp Optom* 1997;80:165-172.

159. Marsh WR, Rawlings SC, Mumma JV. Evaluation of clinical stereoacuity test. *Ophthalmology* 1980;87:1265-1272.
160. Cooper J. Stereopsis. En: Eskridge J, Amos J, Barlett J (eds), *Clinical Procedures in Optometry*. Philadelphia: Lippincott; 1991:121-134.
161. Cooper J, Feldman JM, Medlin D. Comparing stereoscopic performance of children using the Titmus, TNO and Randot stereotest. *J Am Optom Assoc* 1979;50:821-825.
162. Reinecke RD, Simons K. A new stereoscopic test for amblyopia screening. *Am J Ophthalmol* 1974;78:714-721.
163. Clarke WN, Noel LP. Stereoacuity testing in the monofixation syndrome. *J Pediatr Ophthalmol Strabismus* 1990;27:161-163.
164. Fawcett SL, Birch EE. Validity of the Titmus and Randot circles tasks in children with known binocular vision disorders. *J AAPOS* 2003;7:333-338.
165. Hinchliffe HA. Clinical evaluation of stereopsis. *Br Orthopt J* 1978;35:46-57.
166. Simons K, Elhatton K. Artifacts in fusion and stereopsis testing based on red/green dichoptic image separation. *J Pediatr Ophthalmol Strabismus* 1994;31:290-297.
167. Heron G, Dholakia S, Collins D, McLaughlin H. Stereoscopic threshold in children and adults. *Am J Optom Physiol Opt* 1985;62:505-515.
168. Saladin JJ. Stereopsis from a performance perspective. *Optom Vis Sci* 2005;82:186-205.

169. Fawcett SL, Birch EE. Interobserver test-retest reliability of the Randot Preschool Stereoacuity Test. *J AAPOS* 2000;4:354-358.
170. Leat SJ, St Pierre J, Hassan-Abadi S, Faubert J. The moving dynamic Random Dot Stereotest: Development, age norms, and comparison with the Frisby, Randot, and Stereo Smile tests. *J Pediatr Ophthalmol Strabismus* 2001;38:284-294.
171. Tyler CW. Sensory processing of binocular disparity. En: Shor CM, Ciuffreda KJ (eds), *Vergence eye movements; basic and clinical aspects*. Boston: Butterworths; 1983.
172. Hall C. The relationship between clinical stereotest. *Ophthalmol Physiol Opt* 1982;2:133-143.
173. Rosner J, Clift GD. The validity of the Frisby stereotest as a measure of precise stereoacuity. *J Am Optom Assoc* 1984;55:505-506.
174. British Standards Institution. Precision of test methods 1: Guide for the determination and reproducibility for a standard test method. BS 5479. London: HMSO; 1979.
175. Bland JM, Altman DG. Measurement in medicine: The analysis of method comparison studies. *The Statistician* 1983;32:307-317.
176. Bland JM, Altman DG. Comparing two methods of clinical measurement: a personal history. *Int J Epidemiol* 1995;249(suppl.1):S7-S14.
177. Prieto L, Lamarca R, Casado A. La evaluación de la fiabilidad en las observaciones clínicas: el coeficiente de correlación intraclase. *Med Clin (Barc)* 1998;110:142-145.

178. Fleiss JL. The design and analysis of clinical experiments. New York: Wiley; 1986.

179. Macia MA, Lubin P, Rubio P. Psicología matemática II. Madrid: UNED; 1997:130-144.

Anexos

Anexo I. Consentimiento informado



ESCUELA UNIVERSITARIA DE ÓPTICA (UNIVERSIDAD COMPLUTENSE)

Arcos de Jalón s/n 28037 Madrid
Tfno. 91 394 68 94 Fax 91 394 68 85

CONSENTIMIENTO INFORMADO

Fiabilidad y Concordancia de Pruebas de Visión Binocular

Los problemas de visión binocular pueden ser de diversa índole, desde los más conocidos como son la ambliopía u ojo bajo y el estrabismo hasta problemas de tipo funcional no menos frecuentes, pero sí menos conocidos por la población general y cuyos síntomas típicos asociados son dolor ocular, cefaleas, falta de concentración en tareas de cerca, borrosidad, etc.

Con este estudio se pretende estudiar la fiabilidad y la concordancia de diversas pruebas clínicas utilizadas habitualmente para evaluar la visión binocular. Para ello a los participantes en el estudio de investigación se les realizará en dos ocasiones diferentes separadas temporalmente aproximadamente una semana un examen optométrico especializado en pruebas que evalúan el grado de coordinación visual que existe entre ambos ojos.

Debido a la frecuencia con que se presentan estas alteraciones visuales en la población es importante contar con unos métodos de medición de habilidades binoculares que sean suficientemente precisos y fiables. Para poder estudiar cómo de confiables son las diversas pruebas y el grado de concordancia entre diferentes técnicas que evalúan la misma capacidad, es necesario medir a los mismos sujetos en dos ocasiones.

Las personas que conforman la muestra de sujetos del estudio pertenecen al rango de edad donde son frecuentes las alteraciones funcionales que cursan con cansancio asociado a las tareas de cerca (lectura y escritura). De modo que por añadidura tras el análisis de los datos de cada participante nos comprometemos a emitir un informe sobre el estado visual a todos aquellos sujetos que habiendo completado las dos sesiones de pruebas nos lo soliciten. No teniendo derecho a dicho informe aquellos sujetos que sin causa justificada abandonen el estudio tras la primera tanda de medidas.

El procedimiento a seguir se divide en las siguientes fases:

PRIMERA SESIÓN DE EVALUACIÓN:

1. Recogida del cuestionario relativo a la historia visual y ocular del paciente (documento adjunto) y ampliación, si es necesario, de la historia clínica del paciente mediante la formulación de alguna pregunta adicional.
2. Toma de agudeza visual y graduación del paciente.
3. Realización de pruebas de evaluación del sistema motor (acomodación, convergencia) y del sistema sensorial (fusión, estereopsis).

SEGUNDA SESIÓN DE EVALUACIÓN:

4. Realización de pruebas de evaluación del sistema motor (acomodación, convergencia) y del sistema sensorial (fusión, estereopsis).

Antes de comenzar las medidas se dedicará un corto periodo de prueba para que el paciente se familiarice con su cometido en cada prueba. Tras este paso previo, se procederá a la medición.

Si el participante es **usuario de lentes de contacto es necesario que permanezca sin usarlas al menos 24 horas antes de acudir a cada una de las sesiones de evaluación.** De este modo se paliará el posible moldeamiento corneal debido a la presión que ejercen las lentes de contacto y que falsearía los datos de su graduación.

El tiempo total invertido en las medidas no suele superar una hora en la primera sesión y tres cuartos de hora en la segunda. Para la toma de dichas medidas no será necesario el uso de ningún colirio.

La participación en este estudio de investigación no generará ningún problema en su visión, no obstante, no dude en indicarnos cualquier incomodidad que note durante las pruebas o cualquier duda que tenga sobre el experimento. Las medidas podrán ser interrumpidas en cualquier momento si fuera necesario.

D.

.....

con DNI manifiesta que ha sido informado y ha entendido lo referente a su participación en el estudio Fiabilidad y Concordancia de Pruebas de Visión Binocular.

Madrid, dede 2005

Firma del paciente

Firma del responsable de la medida

Beatriz Antona Peñalba
Prfª Titular de la Universidad Complutense de Madrid

Anexo II. Fichas de registro de los resultados de las pruebas

A continuación se incluyen las fichas de anotación utilizadas durante el estudio. Este anexo es un complemento al apartado Metodología General al que el lector puede recurrir para obtener información detallada del uso de estas fichas.

En primer lugar aparece un formulario en el que se recopilaban algunos **datos personales** e información relativa a la **historia clínica** del sujeto participante. Este formulario, junto con el consentimiento informado, era entregado y explicado al sujeto el día de captación de participantes. Ambos documentos debía traerlos cumplimentados el día en que se le citaba para realizar las pruebas **preliminares**, cuya ficha aparece tras el formulario de la historia clínica.

Con toda la información precedente se estaba en disposición de confirmar que el sujeto cumplía los criterios de inclusión en el estudio^{§§§} y se pedía a cada sujeto que pasara a la **extracción las bolas** de una bolsa opaca cuyo color determinaba aleatoriamente el orden de las pruebas que le serían aplicadas en las dos sesiones de medidas. Para facilitar el proceso de peregrinación entre las distintas estaciones, antes de la primera sesión de medidas, se señalaba en las fichas de las seis estaciones (gabinetes) a qué gabinete debía cada examinador enviar al sujeto tras terminar de aplicarle las pruebas correspondientes a su estación; también se anotaba el orden en el que debían aplicarse las pruebas del Gabinete 5 y 6.

^{§§§} Para confirmar que los participantes no presentaban una anomalía acomodativa o de vergencias un optometrista diferente a los que se encargaban de aplicar las pruebas estudiadas se encargaba de analizar en conjunto los datos de la primera sesión de medidas.

Las siguientes **fichas de gabinete** corresponden a las hojas de anotación de las 6 estaciones por las que pasaba el sujeto en cada una de las sesiones de medida identificadas por: Gabinete 1, Gabinete 2, Gabinete 3, Gabinete 4, Gabinete 5 y Gabinete 6. Sólo se han incluido las fichas correspondientes a la 1ª sesión, siendo las correspondientes a la 2ª sesión totalmente equivalentes a éstas. Para evitar que un examinador conociera los resultados recopilados por otros examinadores, cada examinador se encargaba de guardar las fichas de su gabinete, que eran almacenadas sin procesar sus datos al final de cada día de medidas para que el examinador tampoco dispusiera de esa información al repetir las medidas en la segunda sesión.


ESCUELA UNIVERSITARIA DE ÓPTICA (UNIVERSIDAD COMPLUTENSE)

Arcos de Jalón s/n 28037 Madrid
Tfno. 91 394 68 94 Fax 91 394 68 85

Nombre:

Teléfono Fijo:.....Teléfono móvil:.....

CURSO: PRIMERO A B C D (rodea con un círculo)

Recordamos que los sujetos participantes en este estudio no deben estar simultáneamente matriculados o tener aprobadas las siguientes asignaturas del segundo curso: Optometría I, Fundamentos de Contactología, Óptica Oftálmica I.

HISTORIA CLÍNICA

(marque con una X en los cuadrados que corresponda)

SINTOMAS

- | | |
|--|---|
| <input type="checkbox"/> Borrosidad de lejos | <input type="checkbox"/> Dolor de cabeza |
| <input type="checkbox"/> Borrosidad de cerca | <input type="checkbox"/> Visión doble ocasional |
| <input type="checkbox"/> Cansancio visual | <input type="checkbox"/> Sin síntomas visuales |
| <input type="checkbox"/> Irritación ocular | |

HISTORIA OCULAR

- | | |
|---|--|
| <input type="checkbox"/> Ojo vago | <input type="checkbox"/> Traumatismos oculares |
| <input type="checkbox"/> Estrabismo | <input type="checkbox"/> Patología ocular |
| <input type="checkbox"/> Cirugía refractiva | <input type="checkbox"/> Otros:..... |

HISTORIA MÉDICA

- | | |
|--|--------------------------------------|
| <input type="checkbox"/> Diabetes | <input type="checkbox"/> Migrañas |
| <input type="checkbox"/> Hipertensión | <input type="checkbox"/> Otros:..... |
| <input type="checkbox"/> Problemas de tiroides | |

ANTECEDENTES FAMILIARES

- | | |
|-----------------------------------|---------------------------------------|
| <input type="checkbox"/> Glaucoma | <input type="checkbox"/> Hipertensión |
| <input type="checkbox"/> Diabetes | <input type="checkbox"/> Otros:..... |

CORRECCIÓN ÓPTICA ACTUAL

- | | |
|---|----------------------------------|
| <input type="checkbox"/> Gafas | <input type="checkbox"/> Ninguna |
| <input type="checkbox"/> Lentes de contacto | |

Nombre:	Nº
---------	----

PRELIMINARES			DI:	cm
Auto-Rx	OI		AVsc	Avsc AO
	OD		AVsc	
Gafa	OD		AVcc	AVcc AO
	OI		AVcc	
Rx subjetiva	OD		AV	AV AO
	OI		AV	

EXTRACCIÓN DE BOLAS						
	ROJO =	GAB 2	AA alejamiento	Auto-Rx	Randot	Von Graefe
	VERDE =	GAB 3	AA acercamiento	MEM	TNO	Maddox
	AZUL =	GAB 4	AA Lentes -	CCF	Titmus	Thorington

1ª SESIÓN GABINETE 1				Nº	
Retinoscopia Nott OD: _____ cm					
Cover test	VL	Punto neutro: _____	Inversión: _____		
	VP	Punto neutro: _____	Inversión: _____		
Pasar a gabinete		2	3	4	5 ó 6

1ª SESIÓN GABINETE 2				Nº	
AA alejamiento: _____ cm		Autorrefractómetro VP			
FA Bino: _____ cpm					
FA OD: _____ cpm					
FV: _____ cpm					
Pasar a gabinete		2	3	4	5 ó 6

1ª SESIÓN GABINETE 3				Nº					
Retinoscopia MEM OD: _____ D									
AA acercamiento OD: _____ cm		PPC: / _____ cm							
VL	VFN	/	/	Δ	VP	VFN	/	/	Δ
barra	VFP	/	/	Δ	barra	VFP	/	/	Δ
Pasar a gabinete		2	3	4	5 ó 6				

1ª SESIÓN GABINETE 4				Nº					
CCF:		D		AA lentes OD: D					
ARN:		D		ARP: D					
VL	VFN	/	/	Δ	VP	VFN	/	/	Δ
diasp	VFP	/	/	Δ	diasp	VFP	/	/	Δ

Pasar a gabinete	2	3	4	5 ó 6
------------------	---	---	---	-------

1ª SESIÓN GABINETE 5			Nº	
Frisby:		seg arc		
	TNO:	seg arco		
	Randot:	seg arco		
	Titmus:	seg arco		

Pasar a gabinete	6	Dar cita
------------------	---	----------

1ª SESIÓN GABINETE 6				Nº	
Foria Horizontal VL		Foria vertical VP		Foria horizontal VP	
	Von Graefe:		Von Graefe:		Von Graefe:
	Maddox:		Maddox:		Maddox:
			Thorington:		Thorington:

Pasar a gabinete	5	Dar cita
------------------	---	----------

Anexo III. Informe entregado a los participantes sobre su estado visual



ESCUELA UNIVERSITARIA DE ÓPTICA (UNIVERSIDAD COMPLUTENSE)

Arcos de Jalón s/n 28037 Madrid
Tfno. 91 394 68 94 Fax 91 394 68 85

INFORME SOBRE EL ESTADO VISUAL

NOMBRE 1º APELLIDO 2º APELLIDO (Nº 1) participó en el estudio de investigación titulado: "Fiabilidad y Concordancia de Pruebas de Visión Binocular" coordinado por la profesora Beatriz Antona y llevado a cabo durante el curso académico 2004/05. Fruto de este estudio se extrajeron los siguientes datos clínicos:

Estado refractivo

- **Agudeza visual bruta en visión lejana:**
OD: 1,2 OI: 1,5 AO: 1,5
- **Refracción de lejos:**
OD: +0,50 esf -0,00 cil AV: 1,2 AO: 1,5
OI: +0,25 esf 0,00 cil AV: 1,5

Visión binocular

- **Estereoagudeza:**
Frisby: 20 seg arco Randot: 27,5 seg arc
TNO: 30 seg arco Titmus: 40 seg arc
- **Cover test ⁽¹⁾:** VL: -4,5 Δ VP: -3,5 Δ
- **PPC:** 0 / 0 cm
- **Amplitud de Acomodación:**
Alejamiento: 13,39 D
Acercamiento: 16,67 D
Lentes negativas: 8,00 D

- **Flexibilidad Acomodativa Binocular ± 2.00 D:** 8,0 cpm
- **Acomodación Relativa:** ARN: +2,13 D ARP: -2,00 D
- **Respuesta acomodativa:**
Retinoscopía NOTT: 0,25 D
Retinoscopía MEM: 0,50 D
CCF: 0 D
- **Forias horizontales ⁽¹⁾:**
Von Graefe: VP: 0 Δ VL: 0 Δ
Maddox: VP: -0,5 Δ VL: -0,5 Δ
Thorington: VP: 0 Δ
- **Forias verticales ⁽²⁾:**
Von Graefe: VP: 0 Δ
Maddox: VP: 0 Δ
Thorington: VP: 0 Δ
- **Vergencias fusionales horizontales:**
Barra de prismas: VL: VFN: 10,0 / 8,0 Δ
VFP: 40,0 / 15 Δ
VP: VFN: 18 / 14,0 Δ
VFP: 32,5 / 17,0 Δ
Diasporámetro: VL: VFN: 7,5 / 6,5 Δ
VFP: 32 / 15 Δ
VP: VFN: 18 / 12,5 Δ
VFP: 35 / 28 Δ
- **Flexibilidad de vergencias 12 Δ BE / 3 Δ BI:** 10,8 cpm

(1) Un valor negativo corresponde a exoforia y un valor positivo a endoforia.

(2) Un valor negativo corresponde a hiperforia de OI y un valor positivo a hiperforia de OD.

APROVECHAMOS LA OCASIÓN PARA AGRADECERLE SU PARTICIPACIÓN EN EL ESTUDIO Y LE RECOMENDAMOS QUE REVISE SU ESTADO VISUAL UNA VEZ AL AÑO

Anexo IV. Método de Bland y Altman

Este método general de definir el error de medida ha sido recomendado por la *British Standards Institution*¹⁷⁴ y fue descrito por Bland y Altman^{28, 175}. Desde el punto de vista clínico este método tiene la ventaja de que la repetibilidad de una prueba se expresa en las mismas unidades de medida del test y permite al clínico establecer su propio criterio para considerar si un cambio es significativo o no. De forma similar, la diferencia respecto a la media de las diferencias de dos métodos diferentes de medida define el grado de concordancia entre estos dos métodos.

El método de los límites de acuerdo de Bland y Altman²⁸ se aplica a dos tipos de estudios:

- Estimación del acuerdo entre dos medidas de un mismo método (=repetibilidad): Nos preguntamos como de variables son las medidas de un sujeto.
- Estimación de la concordancia o acuerdo entre dos métodos de medida de la misma cualidad visual: Nos preguntamos si dos métodos de medida son intercambiables, es decir, si se puede ignorar con qué método se ha hecho la medida.

A continuación se ampliará la descripción general del método de Bland y Altman en un estudio de repetibilidad, siendo la descripción equivalente en un estudio de concordancia con sólo sustituir las medidas inicial y final por las medidas para cada sujeto obtenidas con una y otra técnica. Por último se incluirá un ejemplo de aplicación del método en un estudio de concordancia entre métodos.

ESTUDIOS DE REPETIBILIDAD

Para valorar la repetibilidad se calcula la diferencia numérica entre las medidas obtenidas en dos diferentes ocasiones y se lleva a cabo un análisis de la distribución de las diferencias entre las dos visitas (inicial y final), extrayendo los límites de acuerdo al 95% para esta distribución⁵. En este análisis, se calcula estadísticamente la diferencia media (DM) como la media de las diferencias, la desviación estándar de las diferencias (DE) y se compara la DM con cero (test-t). La media de la distribución de diferencias será cero si como promedio no hay ninguna diferencia hacia mayor o menor grado al re-evaluar a los mismos sujetos. Los límites del intervalo de acuerdo al 95% acotan la porción central de la distribución de diferencias donde se incluyen el 95% de las diferencias encontradas. Este intervalo de probabilidad se calcula asumiendo simetría en la distribución y sus límites son: $DM \pm 1,96 \times DE$; la anchura de este intervalo se conoce como coeficiente de repetibilidad (CDR).

Para poder aplicar el método de los límites de concordancia al 95% es necesario que se cumplan varios supuestos referentes a los datos:

- Que la media y la desviación estándar de las diferencias sean constantes en todo el rango de medidas. La magnitud de la diferencia no debe estar correlacionada con el rango de datos. Si ocurre esto habrá que estimar los límites de acuerdo por tramos correspondientes a cada rango de valores.
- Que las diferencias se distribuyan aproximadamente según una distribución normal. Para comprobar que la distribución es normal se puede evaluar el histograma de las diferencia o se pueden aplicar pruebas de normalidad como la de Anderson-Darling. Es razonable suponer que con un número suficiente de sujetos la distribución de las diferencias sea normal ya que las diferencias entre dos grupos de datos altamente correlacionados generalmente siguen una

distribución normal incluso si las series de datos originales no se distribuyen normalmente¹⁸.

Un estudio de concordancia es una cuestión de estimación, no un test de hipótesis. Una estimación generalmente conlleva cierto error de muestreo y el modo de estimar los límites de acuerdo explicado anteriormente no es ninguna excepción. Los límites de acuerdo son sólo una estimación de los valores aplicables a la población general y una muestra distinta posiblemente daría límites diferentes.

ESTUDIOS DE CONCORDANCIA ENTRE MÉTODOS. EJEMPLO

A continuación se incluye un ejemplo de un estudio de concordancia para terminar de aclarar cómo se interpretan los parámetros extraídos de este tipo de estudios.

Dados dos métodos de medida, A e B, que sirven para evaluar la misma cualidad y utilizan las mismas unidades. La Tabla 30 muestra los resultados de la medición de dicha cualidad utilizando los dos, así como las diferencias entre ambas medidas para cada sujeto y las diferencias en valor absoluto^{****}.

Se tienen los resultados obtenidos al aplicar cada método en cada uno de los sujetos de la muestra, i : A_i , B_i , y sean las diferencias entre métodos para cada sujeto ($B_i - A_i$). Se calculará el valor medio de las diferencias, $DM = 2,1$ unidades y la desviación estándar de las diferencias, $DE = 37,7$ unidades. De modo que los límites de acuerdo al 95% serán: límite superior: $2,1 - 1,96 \times 37,7 = -73,9$ unidades y límite inferior: $2,1 + 1,96 \times 37,7 = 78,1$ unidades. Por tanto, se espera que una medida hecha con el método B debería estar entre 78,1 unidades más y 73,9 unidades menos que el resultado de una medida tomada

^{****} Datos tomados de: Bland JM, Altman DG. Statistical methods for assessing agreements between two methods of clinical measurement. The Lancet 1986; 307-10

con el método A. La anchura del intervalo entre los límites de concordancia al 95% será 152,0 unidades ó $\pm 76,0$ unidades.

Tabla 30. Resultados de la medición de cierta cualidad utilizando dos métodos A y B (ejemplo).

Sujeto	Método A	Método B	B - A	Valor absoluto de (B - A)
1	494	512	18	18
2	395	430	35	35
3	516	520	4	4
4	434	428	-6	6
5	476	500	24	24
6	557	600	43	43
7	413	364	-49	49
8	442	380	-62	62
9	650	658	8	8
10	433	445	12	12
11	417	432	15	15
12	656	626	-30	30
13	267	260	-7	7
14	478	477	-1	1
15	178	259	81	81
16	423	350	-73	73
17	427	451	24	24

Un método será tanto más repetible cuanto menor sea la diferencia media y el coeficiente de repetibilidad, CDR (intervalo entre los límites de acuerdo al 95%) de la distribución de las diferencias entre medidas repetidas. Igualmente dos métodos de medida de la misma cualidad serán intercambiables o presentarán un buen coeficiente de concordancia (CDC), si la diferencia media y el intervalo entre los límites de acuerdo al 95% de la distribución de las diferencias son bajos respecto a los valores normales del parámetro medido. Los valores obtenidos para el intervalo de concordancia se podrán comparar con los límites que se hayan establecido subjetivamente como clínicamente tolerables al inicio del estudio y se podrá así concluir si los límites de acuerdo obtenidos del análisis de Bland y Altman son o no clínicamente relevantes.

Para visualizar gráficamente los datos es útil utilizar las llamadas gráficas de Bland y Altman^ε (ver Figura 47). El eje de abscisas corresponde al valor medio de ambos métodos para cada sujeto, $(A_i + B_i)/2$. El eje de ordenadas corresponde a la diferencia entre ambos métodos para cada sujeto, $(B_i - A_i)$. En la gráfica correspondiente al ejemplo propuesto se puede apreciar que la diferencia media es muy baja (línea discontinua) y se confunde prácticamente con la recta continua gris correspondiente a una diferencia media de cero. Los límites de acuerdo al 95%, superior e inferior, se muestran con líneas continuas de color negro. En el ejemplo representado en la Figura 47 se puede ver que muy pocas mediciones han concordado (diferencia igual a cero).

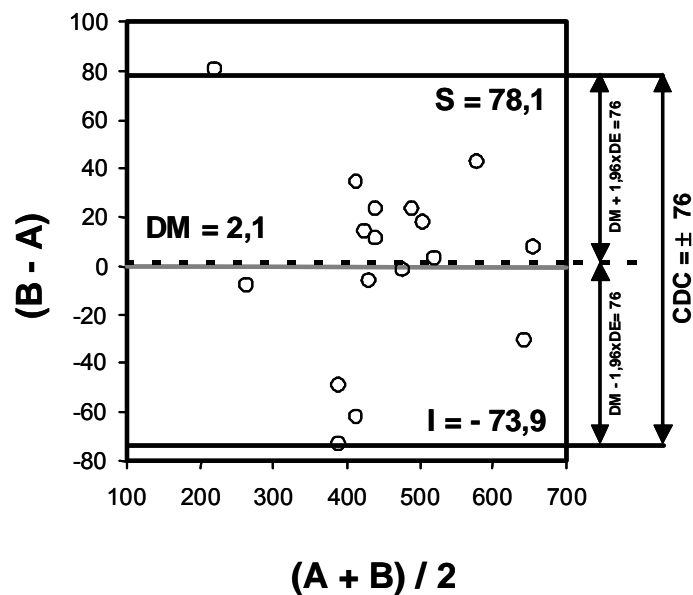


Figura 47. Gráfica de Bland y Altman correspondiente al estudio de concordancia entre los métodos de medida A y B (ejemplo).

^ε Al interpretar visualmente los datos se debe tener en cuenta que en algunos estudios puede haber muchos puntos solapados en el gráfico.

Anexo V. Comparación de métodos de medida

En este apartado se comentarán algunos aspectos importantes que se deben tener en cuenta a la hora de seleccionar el tipo de análisis estadístico adecuado para comparar varios métodos de medida de variables continuas^{†††}. Entre los sistemas de comparación están los siguientes:

ESTUDIOS DE CORRELACIÓN ENTRE PRUEBAS

GRÁFICA DE CORRELACIÓN

Se hace la gráfica con los resultados de cada técnica de medición en uno de los ejes. Como se discutirá más adelante, erróneamente se ha planteado que hay concordancia perfecta entre los dos métodos evaluados, A y B, los puntos correspondientes a los distintos sujetos evaluados (A_i , B_i) se dispondrán en una línea perfecta (ver Figura 48).

COEFICIENTE DE CORRELACIÓN DE PEARSON O SIMPLEMENTE COEFICIENTE DE CORRELACIÓN (r)

Se utiliza para cuantificar el grado de relación lineal entre los dos métodos de medida A y B. En este estudio la hipótesis nula es: “*no existe relación lineal entre los dos métodos*”. El coeficiente de correlación de Pearson (r) se calcula mediante la siguiente expresión:

$$r = \frac{DE_A^2 + DE_B^2 - DE^2}{2DE_A DE_B}$$

DE_A = Desviación estándar del método A

DE_B = Desviación estándar del método B

DE = Desviación estándar de las diferencias entre los métodos A y B

^{†††} Aunque la discusión de este anexo se centra en la comparación de las mediciones realizadas con distintas técnicas de medida; por extensión el razonamiento sería aplicable también a cualquier estudio repetibilidad.

A la vista de la fórmula, para calcular el coeficiente de correlación se puede ver que la existencia de una alta correlación entre dos métodos de medida A y B, no significa que los dos métodos tengan una alta concordancia. Dicha fórmula no incluye la diferencia promedio entre los dos métodos, por lo tanto, la correlación podría ser alta (r cercano a 1) a pesar de que haya diferencias sistemáticas entre los dos métodos.

Siguiendo con el ejemplo propuesto en el Anexo IV. La Figura 48 corresponde a la gráfica de correlación de los métodos de medida A y B. La línea continua gris corresponde a la recta $X=Y$, y la línea continua de color negro es la representación de la recta correspondiente al ajuste lineal cuya expresión también se muestra en la gráfica ($Y = a + b X$). También se incluye el valor del coeficiente de correlación (r).

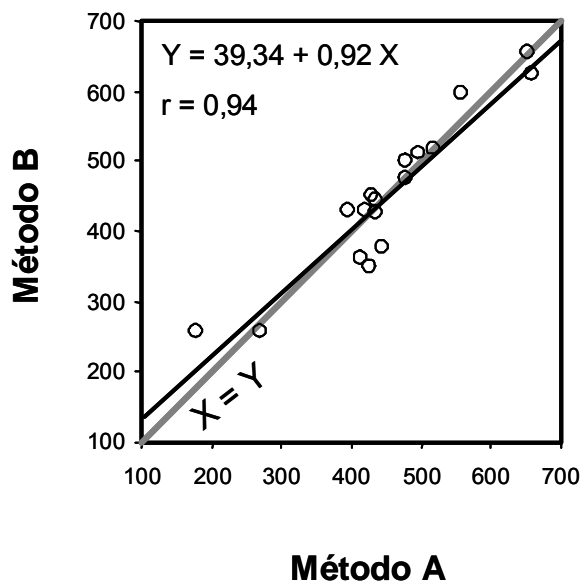


Figura 48. Gráfica de correlación entre los métodos de medida A y B (ejemplo)

El coeficiente de correlación mide si los valores de los dos métodos A y B se disponen en una línea recta de cualquier tipo.

$$Y = a + bX$$

a = Corte con el eje-y

b = Pendiente de la recta

El coeficiente de correlación se limita a decir si dos técnicas están relacionadas linealmente y es de esperar que dos técnicas diseñadas para medir la misma característica lo estén. La idea de que “*si el coeficiente de correlación entre dos medidas es significativamente distinto de cero, la repetibilidad es buena*”, es incorrecta ya que r puede ser muy pequeño y resultar significativo si el tamaño de la muestra es suficientemente grande.

Otras **limitaciones asociadas al uso de r** son:

- r depende del rango de valores observado en la muestra^{28, 176}, es decir, r depende de la variabilidad de los sujetos evaluados. Si la variabilidad entre los sujetos evaluados es pequeña comparada con la variabilidad de las técnicas de medición, r tenderá a ser bajo a pesar de que haya una buena concordancia entre los métodos. Si la variabilidad entre sujetos es grande, la correlación será mayor que si es estrecha. De modo que si se dividiera el rango total en dos partes, cada parte tendría un r menor que el correspondiente al rango total. Por ello es frecuente que los investigadores tiendan a comparar dos métodos de medida sobre el rango total de valores típicamente encontrados, de modo que una alta correlación estará casi garantizada.

- Un cambio en la escala de medidas no afecta a la correlación, pero sí afecta a la concordancia. Si por ejemplo representamos la medida vs. la mitad de la medida, $r=1$, $Y = 2X$; es decir, la correlación es perfecta, pero las medidas claramente no concuerdan, una es dos veces la otra.

Un análisis de correlación es útil cuando se comparan pruebas que no se miden en las mismas unidades.

PRUEBA T PAREADA (t)

La hipótesis en este test es que el promedio de la diferencia entre los dos métodos A y B es cero. Estima la concordancia de los dos métodos en promedio, pero no da información del acuerdo entre los dos métodos para cada uno de los individuos¹⁷⁶. El estadístico t se calcula mediante la siguiente expresión:

$$t = \frac{d - 0}{DE / \sqrt{N}}$$

d = Promedio de las diferencias entre los dos métodos
DE = Desviación estándar de las diferencias entre los dos métodos
N = Número de sujetos de la muestra

Los resultados del test-t también pueden llevar a conclusiones equivocadas independientemente de que exista o no concordancia entre métodos:

- Si se produce una diferencia sistemática entre métodos su concordancia será baja. La DE de las diferencias (en el denominador) será grande y los resultados del test-t tendrán menos probabilidad de ser significativos.
- Si por el contrario, los dos métodos concuerdan en promedio (d tenderá a ser 0) y el test-t será igualmente no significativa.

Es decir, se puede llegar a las mismas conclusiones del test-t tanto si los métodos concuerdan en promedio como si no lo hacen.

COEFICIENTE DE CORRELACIÓN INTRACLASE (CCI)

Desde el punto de vista matemático, el índice más apropiado para cuantificar la concordancia entre diferentes mediciones de una variable numérica es el llamado coeficiente de correlación intraclass (CCI). El CCI estima el promedio de las correlaciones entre todas las posibles ordenaciones por pares de observaciones disponibles y, por tanto, evita el problema de la dependencia del orden del coeficiente de correlación (r). Además, extiende su uso al caso en el que se dispongan de más de dos observaciones por sujeto

Una de las principales limitaciones del CCI es la dificultad de su cálculo, ya que su estimación se realiza de distinta forma dependiendo del diseño del estudio. La forma de cálculo más habitual se basa en un modelo de análisis de varianza (ANOVA) con medidas repetidas¹⁷⁷. La idea es que la variabilidad total de las mediciones se puede descomponer en dos componentes: la variabilidad debida a las diferencias entre los distintos sujetos de la muestra y la variabilidad debida a las diferencias entre las medidas para cada sujeto. Esta última, a su vez, se compone de la variabilidad entre observaciones (debida a diferencias entre los observadores, entre distintas sesiones de medida, entre distintos métodos de medida) y de una variabilidad residual o aleatoria asociada al error que conlleva toda medición. El CCI podría definirse, entonces, como la proporción de la variabilidad total que se debe a la variabilidad de los sujetos. Como toda proporción, el CCI puede tomar valores entre 0 y 1.

CCI = 1 Toda la variabilidad observada se explicaría por diferencias entre sujetos y no por las diferencias entre los métodos de medición o los diferentes observadores. Indica concordancia absoluta entre las medidas.

CCI = 0 Toda la variabilidad observada es debida sólo al error aleatorio asociado a toda medición. Indica ausencia de concordancia entre las medidas.

Aunque toda clasificación de los valores que puede tomar el CCI es subjetiva, resulta útil contar con ella a la hora de interpretar los resultados, por ejemplo se puede utilizar la tabla propuesta por Fleiss¹⁷⁸ (ver Tabla 31).

Tabla 31. Valoración de la concordancia según los valores del coeficiente de correlación intraclass (CCI)

Valor del CCI	Fuerza de la concordancia
> 0,90	Muy buena
0,71 – 0,90	Buena
0,51 – 0,70	Moderada
0,31 – 0,50	Mediocre
< 0,30	Mala o nula

A pesar de que, como se ha comentado, el CCI es el índice más adecuado para cuantificar la concordancia entre diferentes mediciones para el caso de variables numéricas, su uso en la literatura médica, y en concreto en la optométrica⁺⁺⁺, es bajo debido a que presenta varias **limitaciones**:

- Dificultad inherente a su cálculo matemático. Aunque actualmente ya no supone una limitación tan importante pues se han propuesto fórmulas para el cálculo matemático manual del CCI para el caso más sencillo de dos medidas por cada sujeto y también puede calcularse a partir de una tabla ANOVA para medidas repetidas de un factor, ya que este tipo de diseños permite aislar la variabilidad entre sujetos y concentrarnos en la variabilidad debida al factor de interés¹⁷⁷.

⁺⁺⁺ Únicamente un artículo de todos los comentados en los apartados de discusión, ha utilizado este tipo de análisis estadístico (Rouse MW, Borsting E, Deland PN. (2002). Reliability of binocular vision measurements used in the classification of convergence insufficiency. *Optom Vis Sci*, 79: 254-264).

Además, actualmente algunos programas informáticos como el SPSS permiten obtenerlo de forma directa.

- Al tratarse de una prueba paramétrica sólo puede aplicarse a los casos en los que se verifiquen las hipótesis necesarias (variables distribuidas según una normal, igualdad de varianzas e independencia entre errores de cada observador).
- Al igual que el coeficiente de correlación de Pearson, el valor del CCI depende en gran medida de la variabilidad de los valores observados. Cuanto más homogénea sea la muestra estudiada (los sujetos varían poco en sus puntuaciones), más bajo tenderá a ser el valor del CCI. Si la muestra es heterogénea, el CCI tiende a ser mayor.
- Como en muchos otros índices estadísticos, los resultados obtenidos para el CCI están expresados en términos absolutos y carece de interpretación clínica. Por ejemplo diferencias sistemáticas de 5 unidades en la medida de longitudes para todas las observaciones proporcionan el mismo CCI tanto si realizamos las mediciones en metros, como si se realizan en kilómetros. Por tanto el investigador debe tener en cuenta la significación clínica que de las diferencias observadas tienen para su estudio. Aunque el CCI es un índice objetivo que puede facilitar la comparación de distintos métodos de medida, la carencia de interpretación clínica, ha propiciado la aparición de otros métodos de análisis mucho más intuitivos y fácilmente interpretables como el método gráfico de Bland y Altman (ver Anexo IV).

MÉTODO GRÁFICO DE BLAND Y ALTMAN

GRÁFICA DE PROMEDIO Y DIFERENCIA DE LAS MEDIDAS

En esta gráfica el eje-x corresponde al promedio de las dos mediciones, $(A + B)/2$, y el eje-y a la diferencia de los valores de los dos métodos, $(B - A)$. Esta gráfica proporciona más información que la gráfica de correlación entre pruebas pues permite evaluar las diferencias en cada individuo. Si los dos métodos generan en cada sujeto aproximadamente el mismo valor, todos los puntos de la gráfica deben estar próximos a $y = 0$. Por el contrario, cuanto más esparcidos estén los puntos respecto a este valor, $y = 0$, más diferentes serán los resultados de ambos métodos.

LÍMITES DE ACUERDO

Estos límites se calculan para estimar el grado de dispersión de las diferencias entre los dos métodos. Si las diferencias entre los dos métodos tienen una distribución normal^ξ y la diferencia de los dos métodos es independiente de la magnitud de la variable medida, se espera que el 95% de las diferencias caigan dentro de $\pm 1,96$ desviaciones estándar de las diferencias. Es habitual reflejar estos límites de acuerdo en la gráfica comentada en el apartado anterior. Este método estadístico se conoce como método de Bland y Altman, que se ha descrito en el Anexo IV.

Si las diferencias que se encuentran en $\pm 1,96$ desviaciones estándar no son clínicamente importantes, los dos métodos de medición pueden ser considerados como intercambiables. Por tanto, que la diferencia sea aceptable o no es una decisión clínica y no estadística. Es decir, depende de si la concordancia encontrada es suficiente para el propósito clínico particular. Si las diferencias tienen relevancia

^ξ Para comprobar que la distribución es normal se puede evaluar el histograma de las diferencias o se pueden aplicar pruebas de normalidad como la de Anderson-Darling.

clínica, un método no puede ser reemplazado por el otro y se dirá que la concordancia entre métodos es baja.

ANÁLISIS DE VARIANZA (ANOVA) DE DOS FACTORES CON MEDIDAS REPETIDAS EN AMBOS

Dada la complejidad de este tipo de análisis, está más allá de los objetivos de este apartado el hacer una descripción detallada del mismo que, por otra parte, puede consultarse en los manuales de estadística¹⁷⁹. La discusión se limitará a comentar algunos aspectos generales relacionados con el diseño de ANOVA utilizado en el presente trabajo.

Para comparar más de 2 métodos se podrían comparar de 2 en 2 por ejemplo con una prueba t, pero además de ser un poco tedioso, no tendría en cuenta la posible interacción entre variables. La técnica a usar es un análisis de varianza. En este tipo de análisis la hipótesis nula general afirma que no existe diferencia alguna entre las medias en la variable dependiente de los distintos grupos de medidas. Las diferencias podrían tener dos fuentes de variabilidad: variabilidad debida a las distintas situaciones experimentales (por ejemplo, método de medida, sesión de medidas) y la debida al error experimental. El ANOVA se fundamenta en el estudio de estas variabilidades.

Los análisis de varianza de dos factores son modelos en los que se quiere estudiar la influencia de dos variables independientes (factores) sobre una variable dependiente. Hay varios diseños posibles por lo que se centrará la discusión en el modelo aplicable al conjunto de datos del estudio presente, es decir, a un ANOVA de dos factores con medidas repetidas en ambos factores¹⁷⁹. En un diseño de estas características existen tres efectos de interés: (1) el efecto individual del primer factor (A = técnica de medida, con T niveles), (2) el efecto individual del segundo factor (B = sesión de medida, con I niveles) y (3) el efecto conjunto de la interacción entre los dos factores.

Las hipótesis de este tipo de modelos son:

- Las T medias del factor A son iguales
- Las I medias del factor B son iguales
- El efecto de la interacción es nulo

En un ANOVA de dos factores, ambos con medidas repetidas, los sujetos que participan en el experimento pasan por todas las condiciones experimentales, es decir, por todas las condiciones definidas por las combinaciones entre los niveles de ambos factores. Eligiendo una muestra de sujetos que pase por todos los niveles de los factores estudiados, se controla la influencia de las diferencias individuales sobre los tratamientos. Sin embargo, estos diseños conllevan el peligro de que la secuencia en que se les pasan los procedimientos influya en los resultados. En algunos casos el efecto de la secuencia se puede controlar contrabalanceando el orden en que se aplican los procedimientos, pero en otros casos la práctica, la fatiga, la transferencia de aprendizaje, etc. pueden ejercer un profundo efecto sobre los resultados, por lo que, en esos casos no se deberían utilizar los diseños de medidas repetidas.

Un ANOVA nos dice si hay diferencias entre los resultados de una variable dependiente en función de los distintos niveles de distintas variables independientes. Lo único que se obtiene son varias razones F que, en caso de que alguna sea significativa (es decir, se rechace la hipótesis correspondiente) nos indicaría que al menos uno de los niveles del factor en cuestión es significativo, es decir, que al menos entre dos medias hay diferencias, pero no nos dice entre qué métodos está esa diferencia significativa. Las comparaciones múltiples permiten establecer una información más exacta sobre la importancia de cada uno de los niveles de la variable o variables independientes como causa de las diferencias. Se deben hacer comparaciones no planificadas, a posteriori o *post-hoc*, siendo algunas de las más

utilizadas: la prueba de Turkey o la prueba de Scheffé, siendo esta última la que se utilizó en este estudio.

Anexo VI. Contribuciones científicas derivadas del desarrollo del doctorado

En este apartado se recogen las publicaciones en revistas científicas y las contribuciones a congresos que hasta el momento han servido para divulgar los resultados derivados del desarrollo del doctorado.

PUBLICACIONES RESULTANTES DEL PERIODO DE INVESTIGACIÓN PARA LA CONCESIÓN DEL DIPLOMA DE ESTUDIOS AVANZADOS (DEA)

- Antona B, Barra F, Barrio AR, González E, Sánchez I. The validity and repeatability of the new aniseikonia test. *Optom Vis Sci*, 2006, 83: 903-909
- Antona B, Barra F, Barrio AR, González E, Sánchez I. Validity and repeatability of a new test of aniseikonia. *IOVS*, 2007, 48: 58-62

PUBLICACIONES RESULTANTES DEL DESARROLLO DE LA TESIS DOCTORAL

- Antona B, Barrio AR, Barra F, González E, Sánchez I. Repeatability and agreement in amplitude of fusional vergences measurements. *Ophthal Physiol Opt*, 2007 (enviada revisión, en espera de su aceptación)
- Antona B, Barrio AR, Barra F, González E, Sánchez I. Repeatability intra-examiner and agreement in amplitude of accommodation measurements. *Vis Res*, 2008 (enviada)
- Antona B, Barrio AR, Barra F, González E, Sánchez I. Repeatability intra-examiner and comparison among methods of measuring accommodative response. *Brit J Ophthalmol*, 2008 (en preparación)

CONGRESOS INTERNACIONALES RESULTANTES DEL PERIODO DE INVESTIGACIÓN PARA LA CONCESIÓN DEL DIPLOMA DE ESTUDIOS AVANZADOS (DEA)

- Antona B, Barra F, Barrio AR. Medida clínica de la aniseiconia. XVIII Congreso Internacional de Optometría e Contactología, organizado por UPOOP. Sintra (Portugal), 2004
- Barra F, Antona B, Barrio AR, González E, Sánchez I. Validity and Repeatability of Two Direct-Comparison Tests of Aniseikonia. ARVO 2005 Fort Lauderdale Florida (EEUU), 2005

CONGRESOS INTERNACIONALES RESULTANTES DEL DESARROLLO DE LA TESIS DOCTORAL

- Antona B, Barra F, González E, Sánchez I, Barrio A. Repetibilidad y acuerdo de la medida de reservas fusionales. 19 Congreso Internacional de Optometría, Contactología y Óptica Oftálmica. Madrid (España), 2006
- Barrio A, Antona B, Barra F, González E, Sánchez I. Repeatability and concordance of the clinical measurement of the amplitude of accommodation. ARVO 2007 Fort Lauderdale, Florida (EEUU), 2007
- Barra F, Antona B, González E, Sánchez I, Barrio A. Repetibilidad y acuerdo de la medida de estereoagudeza. 20 Congreso Internacional de Optometría, Contactología y Óptica Oftálmica. Madrid (España), 2008 (Comunicación Científica Premio del Congreso)
- Antona B, Barra F, González E, Sánchez I, Barrio A, González G. Repetibilidad de la medida de flexibilidad acomodativa. 20 Congreso Internacional de Optometría, Contactología y Óptica Oftálmica. Madrid (España), 2008
- Barra F, Antona B, Barrio A, González E, Sánchez I. Repeatability and agreement in accommodative response measurement. ARVO 2008 Fort Lauderdale, Florida (EEUU), 2008 (aceptado)

Anexo VII. Proyectos de investigación relacionados con el desarrollo del doctorado

En este apartado se recogen los proyectos de investigación que constituyen la fuente de financiación de este trabajo.

Título del proyecto: **COMPARACIÓN CLÍNICA DE DOS TEST DE ANISEICONIA (PR3/04-12368)**

Entidad financiadora: Universidad Complutense de Madrid (UCM)
Entidades participantes: Escuela Universitaria de Óptica (UCM)
Duración, desde: 01-01-2004 hasta: 31-12-2004 Cuantía de la subvención: 2300 €
Investigador responsable: FRANCISCO BARRA LÁZARO
Número de investigadores participantes: 5

Título del proyecto: **FIABILIDAD Y CONCORDANCIA DE PRUEBAS CLÍNICAS EVALUADORAS DEL SISTEMA ACOMODATIVO (PR/07-14909)**

Entidad financiadora: Universidad Complutense de Madrid (UCM)
Entidades participantes: Escuela Universitaria de Óptica (UCM)
Duración, desde: 01-01-2007 hasta: 10-12-2007 Cuantía de la subvención: 3000 €
Investigador responsable: ANA ROSA BARRIO DE SANTOS
Número de investigadores participantes: 5

Título del proyecto: **FIABILIDAD Y CONCORDANCIA DE PRUEBAS CLÍNICAS EVALUADORAS DE LA DESVIACIÓN OCULAR (PR34/07-15876)**

Entidad financiadora: Banco Santander Central Hispano (BSCH)
Entidades participantes: Universidad Complutense de Madrid (UCM)
Duración, desde: 01-01-2008 hasta: 31/12-2009 Cuantía de la subvención: 10000 €
Investigador responsable: FRANCISCO BARRA LÁZARO
Número de investigadores participantes: 5

Nombre de archivo: TESIS BEATRIZ ANTONA - publicacion electronica
Directorio: G:\TESIS\publicacion electronica tesis
Plantilla: C:\Documents and Settings\ucm\Datos de
programa\Microsoft\Plantillas\Normal.dot
Título: I
Asunto:
Autor: Cuculí
Palabras clave:
Comentarios:
Fecha de creación: 03/12/2009 1:05
Cambio número: 2
Guardado el: 03/12/2009 1:05
Guardado por: ucm
Tiempo de edición: 9 minutos
Impreso el: 03/12/2009 1:13
Última impresión completa
Número de páginas: 276
Número de palabras: 97.356 (aprox.)
Número de caracteres:535.460 (aprox.)