

CERVANTES

Batería de Evaluación de la Inteligencia
6 – 18 años

Manual Estadístico

Marzo de 2.009



Grupo
ALBOR-COHS

ÍNDICE DE CONTENIDOS

	Pág.	
1	Introducción	3
2	Construcción de la prueba	4
	2.1. Selección y orden de los ítems	4
	2.2. Distribuciones de puntuaciones	28
3	Muestras de Baremación	39
	3.1. Distribución de las edades	39
	3.2. Distribución geográfica	40
	3.3. Distribución por centros escolares	41
4	Estandarización de las puntuaciones	43
	4.1. Homogeneidad de los perfiles	51
	4.2. Intervalos de confianza para la estimación de la puntuación verdadera	55
5	Propiedades Psicométricas	57
	5.1. Validez de constructo, concurrente y discriminante	57
	5.2. Fiabilidad	60
	5.3. Distribución de las puntuaciones por edad y sexo	71
	5.4. Efectos "suelo" y "techo"	74

1. Introducción

Los análisis estadísticos que se exponen en este Manual tienen un objetivo doble, eminentemente práctico, para los usuarios de la **Batería CERVANTES**. Por una parte, se facilitan para que el profesional disponga de la información necesaria para valorar la utilidad y la precisión de la medida de la capacidad de razonamiento que puede obtener con el empleo de este instrumento. Por otra parte, tanto profesionales como docentes e investigadores pueden conocer los datos que permiten entender el fundamento estadístico de los resultados que proporciona, en forma de perfil e intervalos de confianza.

Los análisis estadísticos y gráficos mostrados se han llevado a cabo empleando el paquete estadístico R versión 2.7.2. Para los análisis de escalas acumulativas de Mokken y el estudio de homogeneidad monótona del conjunto de ítems se ha utilizado la librería "Mokken" de R (L. Andries van der Ark (2007). Mokken: Mokken Scale Analysis in R. R package version 0.1).

Este manual se modificará en el futuro, a medida que el equipo de investigación disponga de nuevos datos sobre la **Batería CERVANTES**. Las modificaciones y ampliaciones se harán públicas en el sitio web: www.bateriacervantes.es,

Es intención del Grupo editor continuar durante los próximos dos años ampliando baremos en todos los grupos de edad, grupos específicos de población y estudios con población adulta.

2. Construcción de la prueba

2.1. Selección y orden de los ítems

2.1.1. Ítems iniciales

Para la elaboración de las 3 pruebas generales de la **Batería CERVANTES**, se diseñaron 140 ítems que cumplieren los siguientes criterios:

1. Los requerimientos culturales para su resolución debían ser mínimos.
2. Para su comprensión y resolución no se debía necesitar ningún tipo de lenguaje oral o escrito.
3. En su ejecución no debía ser necesario un grado de destreza motora fina, (este requisito sí se requiere para la prueba complementaria de "Diseños")

Una vez diseñados los ítems, se llevó a cabo una aplicación piloto con 298 escolares, de edades comprendidas entre los 6 y los 14 años, cuyas respuestas nos permitieron advertir que algunas características de las pruebas, irrelevantes para los adultos, resultaban confusas para ellos, por lo que se procedió a realizar pequeños cambios en algunos dibujos-estímulos.

Para establecer la dificultad de los ítems, éstos se ordenaron de dos maneras diferentes, con las cuales se realizaron varias aplicaciones. Con ello se pretendió evitar la aparición de errores en los ítems presentados en las últimas posiciones, que se debieran más a la fatiga o la frustración por fallos, que a la propia dificultad del ítem; facilitando de este modo que se obtuviera, para cada ítem, un índice de dificultad menos sesgado por la posición ocupada en la prueba.

2.1.2. Metodología de Mokken para la selección y categorización en dificultad de los ítems

En la **Batería CERVANTES** se propone a cada sujeto en evaluación llevar a cabo diversas tareas con diferentes grados de dificultad de ejecución, las cuales permitan situarle en una determinada posición respecto a la población general de su mismo grupo de edad, a partir del nivel alcanzado en la prueba (estimado a partir del número de ítems acertados)

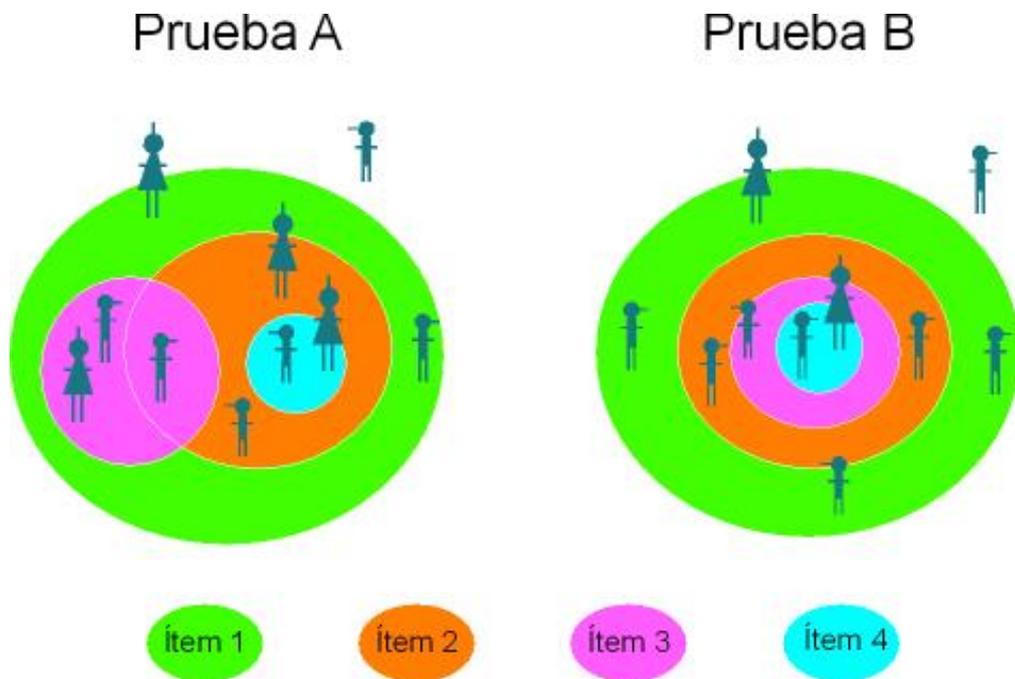
Contando con ello, considerábamos que el proceso de selección de los ítems que formarían el instrumento debería proporcionar un conjunto final en el que:

- La dificultad de los diferentes ítems se distribuyese de forma tal que se evite una excesiva presencia de ítems similares. No es conveniente que haya ni demasiados ítems fáciles ni demasiados ítems difíciles.
- Los ítems se encontrasen ordenados del más fácil al más difícil de resolver. De esta forma es posible acortar el tiempo de las aplicaciones estableciendo un límite de errores.
- El incremento de la dificultad debía ser lo más uniforme posible. Por ejemplo, en una prueba de 10 ítems, el incremento de la dificultad entre cada uno de ellos debería ser cercano al 10%. De esta manera se conseguiría discriminar mejor entre dos sujetos que respondiesen correctamente un número consecutivo de ítems.

Tradicionalmente, se ha establecido el índice de dificultad de un ítem en función del porcentaje de veces que ha sido acertado, creándose las escalas ordenando los ítems por dificultad creciente. Sin embargo, en la elaboración de la **Batería CERVANTES** hemos desestimado este método, ya que, a nuestro entender, conlleva al menos dos deficiencias:

- No ofrece garantías de que los ítems más difíciles tengan relación con el resto de los ítems. Dado el bajo nivel de aciertos de estos ítems, no se puede discriminar si son el resultado de la propia inteligencia lógica o del azar.
- El ítem más difícil no es aquél que menos aciertos obtiene, sino aquél que sólo lo aciertan los individuos que obtienen puntuaciones más altas.

En la siguiente figura se muestra esquemáticamente la necesidad de aplicar una metodología como el Análisis de Escalas de Mokken a la hora de ordenar una serie de ítems en una prueba.



Las dos pruebas constan de 4 ítems de la misma dificultad (calculado como el número de sujetos que dan la respuesta correcta dentro del círculo entre el total de sujetos). La Dificultad del ítem 1 en las dos pruebas es de 9/10, del ítem 2 es de 5/10, del ítem 3 es de 3/10 y del ítem 4 es de 2/10. A priori, ambas pruebas colocarían a un sujeto de la misma manera dentro de la población general respecto de la característica latente que miden ambas.

El problema de la Prueba A reside en el ítem 3, ya que hay personas que acertando un ítem más difícil (ítem 4) no lo aciertan y los que aciertan el ítem 3 no aciertan necesariamente un ítem más fácil (ítem 2).

La situación ideal es la dada por la Prueba B.

¹ Para poder seguir la explicación del gráfico, es necesario imprimir esta página en color.

Por todo ello, estimamos oportuno seleccionar los ítems que conformarían cada escala utilizando el modelo probabilístico de Mokken para ítems dicotómicos (Martínez Arias, M.R. y Rivas Moya, T., 1990). Este modelo establece una serie de criterios para determinar el grado de acumulatividad o escalabilidad de un conjunto de ítems. Cuando se cumplen dichos criterios, se entiende que los ítems se encuentran ordenados en dificultad, conformando una escala uniforme en la que un acierto de uno cualquiera de ellos implica una alta probabilidad de haber acertado los anteriores.

Los criterios que debe cumplir un conjunto de ítems para aceptarse su escalabilidad, según el modelo de Mokken, son:

Homogeneidad de los ítems de la escala (H): indica el grado de asociación entre todos los ítems de la escala. Representa el grado de escalabilidad del conjunto de ítems. Siempre que es superior a 0,30, se acepta la escalabilidad, es decir, los ítems se encuentran ordenados en dificultad y, además, el acierto de un ítem implica una alta probabilidad de acertar otros más fáciles. Se puede interpretar su valor con esta tabla (Rivas Moya, T, 1999):

Tabla 1: Valoración de los Coeficientes H

$0,5 < H$	Escala Fuerte
$0,4 < H < 0,5$	Escala Media
$0,3 < H < 0,4$	Escala Débil

Homogeneidad de un ítem (Hi): indica el grado de asociación de ítem con el resto. Se acepta si es superior a 0,3.

Homogeneidad entre par de ítems (Hij): indica el grado de asociación entre cada par de ítems. Este valor es máximo, 1, cuando todos los que aciertan un ítem, también aciertan los de menor dificultad.

Se acepta si es diferente de 0.

Monotonía de las dificultades de los ítems. Presentadas las matrices P (+,+) (proporción de sujetos que han acertado correctamente cada par de ítems) y P (-,-) (proporción de sujetos que han fallado cada par de ítems), las puntuaciones obtenidas en todos los pares deben formar una escala no decreciente.

2.1.3. Estudio del coeficiente H

De los 53 ítems inicialmente preparados para la creación de la Prueba **Semejanzas**, encontramos una colección de 27 ítems que cumplían las condiciones establecidas por Mokken. De los 55 ítems inicialmente preparados para la creación de la Prueba **Analogías**, encontramos una colección de 39 ítems que cumplían las condiciones establecidas por Mokken. De los 50 ítems inicialmente preparados para la creación de la Prueba **Contextos**, encontramos una colección de 33 ítems que cumplían las condiciones establecidas por Mokken.

Coincidiendo con Snijders, J.TH., Tellegen, P.J. y Laros, J.A., 1989, estimamos oportuno administrar esta escala en una forma adaptada que permitiera:

1. Limitar el número de ítems a aplicar, reduciendo la duración de la aplicación de la prueba.
2. Limitar la administración de ítems demasiado fáciles o difíciles, evitando la posibilidad de fatiga o frustración.

Para ello, se procedió a dividir la escala en series, cuyos ítems eran equivalentes en nivel de dificultad, de forma que se pudiera establecer un procedimiento de administración adaptado equivalente a la administración de la escala completa. De esta forma, se obtuvieron tres nuevas escalas o series:

Tabla 1 Coeficiente de Escalabilidad (H) y Significación (p) de las tres series de la Prueba Semejanzas

Serie	Ítems	H	Δ^*	p
A	9	0,47	53,25	0,00
B	9	0,69	75,64	0,00
C	9	0,72	77,18	0,00

Δ^* Estadístico de contraste

Tabla 2 Coeficiente de Escalabilidad (H) y Significación (p) de las tres series de la Prueba Analogías

Serie	Ítems	H	\square^*	p
A	13	0,75	124,20	0,00
B	13	0,79	121,07	0,00
C	13	0,83	137,50	0,00

Δ^* Estadístico de contraste

Tabla 3 Coeficiente de Escalabilidad (H) y Significación (p) de las tres series de la Prueba Contextos

Serie	Ítems	H	\square^*	p
A	11	0,71	98,61	0,00
B	11	0,72	103,44	0,00
C	11	0,80	103,00	0,00

Δ^* Estadístico de contraste

Según la clasificación sugerida por Mokken todas las series forman escalas fuertes y la Serie A de la Prueba Semejanzas una escala media.

Como indican Martínez Arias, M.R. y Rivas Moya, T., 1990, Mokken también estableció las bases para llevar a cabo el contraste de la hipótesis de aleatoriedad. Es decir, en qué medida esos valores obtenidos en las diferentes escalas son significativamente distintos de cero y, por tanto, no debidos al azar. El estadístico de contraste es Δ^* .

En todos los casos, el valor de Δ^* es mayor que el esperado bajo la hipótesis nula (valores son debidos al azar) por lo que aceptamos, que el coeficiente H_i o coeficiente de escalabilidad de la escala en las 9 series no es debido al azar.

2.1.4. Estudio de los coeficientes H_i y H_{ij}

A continuación se ofrecen las tablas con los coeficientes de homogeneidad de cada ítem (H_i) y de cada par de ítems (H_{ij}). Se observa que todos los H_i son superiores a 0,30 y que todos los H_{ij} son distintos de cero, aceptándose la escalabilidad de las 9 series.

2.1.4.1. Estudio de los coeficientes H_i y H_{ij} de la prueba Semejanzas

Tabla 4 Índices de Homogeneidad de cada par de ítems de la Serie A de la Prueba Semejanzas

	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9
Ítem 1	0,49	0,42	0,63	0,72	0,83	0,93	0,43	0,16
Ítem 2		0,28	0,56	0,84	0,92	0,92	1,00	1,00
Ítem 3			0,06	0,30	0,39	0,53	0,69	1,00
Ítem 4				0,33	0,43	0,49	0,59	0,79
Ítem 5					0,69	0,77	0,86	1,00
Ítem 6						0,74	0,81	1,00
Ítem 7							0,72	1,00
Ítem 8								0,35

Tabla 5 Índices de Homogeneidad de cada Ítem (H_i) y Significación (p) de la Serie A de la Prueba Semejanzas

	H_i	Δ^*	p
Ítem 1	0,61	14,64	0,00
Ítem 2	0,63	18,85	0,00
Ítem 3	0,29	21,02	0,00
Ítem 4	0,32	24,34	0,00
Ítem 5	0,50	37,54	0,00
Ítem 6	0,58	39,16	0,00
Ítem 7	0,67	31,21	0,00
Ítem 8	0,70	17,38	0,00
Ítem 9	0,77	12,26	0,00

Δ^* Estadístico de contraste

En la anterior tabla se muestra el valor estimado H_i , para cada ítem de la Serie A de la prueba Semejanzas. Para cada H_i se han planteado las hipótesis:

Hipótesis Nula: $H_i=0$

Hipótesis Alternativa: $H_i>0$

Para muestras de tamaño suficientemente grande, el estadístico del contraste, Δ^* , se distribuye según una distribución normal de media 0 y desviación típica 1 (Mokken, 1971, pág.163) y consideraremos que H_i es distinto de cero si $\Delta^* \geq Z_{1-\alpha}$

Observamos como en todos los casos rechazamos la hipótesis nula, por lo que $H_i>0$ para todos los ítems de la Serie A de la Prueba Semejanzas.

Tabla 6 Índices de Homogeneidad de cada par de ítems (H_{ij}) de la Serie B de la Prueba Semejanzas

	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9
Ítem 1	0,56	0,66	0,70	0,87	0,89	0,79	0,80	0,80
Ítem 2		0,71	0,70	0,86	0,91	0,82	0,92	1,00
Ítem 3			0,43	0,69	0,82	0,85	0,88	0,87
Ítem 4				0,60	0,73	0,77	0,92	0,98
Ítem 5					0,48	0,66	0,77	0,88
Ítem 6						0,64	0,78	0,86
Ítem 7							0,69	0,82
Ítem 8								0,68

Tabla 7 Índices de Homogeneidad de cada Ítem (H_i) y Significación (p) de la Serie B de la Prueba Semejanzas

	H_i	Δ^*	p
Ítem 1	0,71	20,32	0,00
Ítem 2	0,74	22,45	0,00
Ítem 3	0,67	33,96	0,00
Ítem 4	0,65	35,32	0,00
Ítem 5	0,63	42,18	0,00
Ítem 6	0,66	44,80	0,00
Ítem 7	0,71	44,14	0,00
Ítem 8	0,76	38,50	0,00
Ítem 9	0,81	31,73	0,00

Δ^* Estadístico de contraste

Observamos como en todos los casos rechazamos la hipótesis nula, por lo que $H_i > 0$ para todos los ítems de la Serie B de la Prueba Semejanzas.

Tabla 8 Índices de Homogeneidad de cada par de ítems (H_{ij}) de la Serie C de la Prueba Semejanzas

	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9
Ítem 1	0,72	0,75	0,86	0,90	0,97	0,91	0,92	1,00
Ítem 2		0,54	0,72	0,85	0,90	0,91	0,90	1,00
Ítem 3			0,61	0,76	0,88	0,92	0,98	1,00
Ítem 4				0,60	0,74	0,87	0,91	0,97
Ítem 5					0,52	0,73	0,85	1,00
Ítem 6						0,62	0,77	0,89
Ítem 7							0,55	0,80
Ítem 8								0,60

Tabla 9 Índices de Homogeneidad de cada Ítem (H_i) y Significación (p) de la Serie C de la Prueba Semejanzas

	H_i	Δ^*	p
Ítem 1	0,84	26,84	0,00
Ítem 2	0,76	34,79	0,00
Ítem 3	0,73	36,82	0,00
Ítem 4	0,72	41,65	0,00
Ítem 5	0,68	44,00	0,00
Ítem 6	0,68	44,27	0,00
Ítem 7	0,71	39,80	0,00
Ítem 8	0,72	33,19	0,00
Ítem 9	0,81	24,56	0,00

Δ^* Estadístico de contraste

Observamos como en todos los casos rechazamos la hipótesis nula, por lo que $H_i > 0$ para todos los ítems de la Serie C de la Prueba Semejanzas.

2.1.4.2. Estudio de los coeficientes H_i y H_{ij} de la prueba Analogías

Tabla 10 Índices de Homogeneidad de cada par de ítems (H_{ij}) de la Serie A de la Prueba Analogías

	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9	Ítem 10	Ítem 11	Ítem 12	Ítem 13
Ítem 1	0,72	0,86	0,91	0,95	0,92	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Ítem 2		0,46	0,42	0,39	0,49	0,65	0,75	0,80	0,80	0,97	0,98	0,94
Ítem 3			0,69	0,76	0,83	0,91	0,94	0,97	0,97	0,95	1,00	1,00
Ítem 4				0,68	0,75	0,85	0,95	0,96	0,95	0,95	0,96	1,00
Ítem 5					0,71	0,81	0,87	0,93	0,96	0,96	0,97	1,00
Ítem 6						0,62	0,71	0,79	0,81	0,86	0,91	0,96
Ítem 7							0,73	0,76	0,83	0,89	0,88	0,93
Ítem 8								0,80	0,82	0,91	0,92	0,88
Ítem 9									0,86	0,90	0,93	0,97
Ítem 10										0,67	0,75	0,84
Ítem 11											0,79	0,80
Ítem 12												0,85

Tabla 11 Índices de Homogeneidad de cada ítem (H_i) y Significación (p) de la Serie A de la Prueba Analogías

	H_i	Δ^*	p
Ítem 1	0,90	22,93	0,00
Ítem 2	0,57	45,28	0,00
Ítem 3	0,76	37,90	0,00
Ítem 4	0,73	46,83	0,00
Ítem 5	0,71	54,67	0,00
Ítem 6	0,70	57,09	0,00
Ítem 7	0,75	61,74	0,00
Ítem 8	0,81	63,83	0,00
Ítem 9	0,84	64,34	0,00
Ítem 10	0,82	52,39	0,00
Ítem 11	0,86	47,17	0,00
Ítem 12	0,88	41,38	0,00
Ítem 13	0,90	29,25	0,00

Δ^* Estadístico de contraste

En la anterior tabla se muestra el valor estimado H_i , para cada ítem de la Serie A de la prueba Analogías. Para cada H_i se han planteado las hipótesis:

Hipótesis Nula: $H_i=0$

Hipótesis Alternativa: $H_i>0$

Para muestras de tamaño suficientemente grande, el estadístico del contraste, Δ^* , se distribuye según una distribución normal de media 0 y desviación típica 1 (Mokken, 1971, pág.163) y consideraremos que H_i es distinto de cero si $\Delta^* \geq Z_{1-\alpha}$.

Observamos como en todos los casos rechazamos la hipótesis nula, por lo que $H_i > 0$ para todos los ítems de la Serie A de la Prueba Analogías.

Tabla 12 Índices de Homogeneidad de cada par de ítems (H_{ij}) de la Serie B de la Prueba Analogías

	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9	Ítem 10	Ítem 11	Ítem 12	Ítem 13
Ítem 1	0,79	0,83	0,91	0,93	0,72	0,65	0,74	0,72	0,85	1,00	0,81	1,00
Ítem 2		0,78	0,85	0,85	0,65	0,62	0,75	0,73	0,85	1,00	0,82	1,00
Ítem 3			0,65	0,77	0,78	0,79	0,88	0,87	0,93	1,00	0,91	1,00
Ítem 4				0,64	0,76	0,85	0,94	0,93	0,97	1,00	0,96	1,00
Ítem 5					0,48	0,65	0,85	0,91	0,96	0,97	0,98	1,00
Ítem 6						0,67	0,84	0,95	0,98	1,00	0,98	1,00
Ítem 7							0,75	0,87	0,92	0,95	0,95	1,00
Ítem 8								0,56	0,70	0,80	0,86	0,90
Ítem 9									0,82	0,87	0,91	0,98
Ítem 10										0,80	0,91	0,92
Ítem 11											0,76	0,82
Ítem 12												0,73

Tabla 13 Índices de Homogeneidad de cada Ítem (H_i) y Significación (p) de la Serie B de la Prueba Analogías

	H_i	Δ^*	p
Ítem 1	0,81	32,01	0,00
Ítem 2	0,77	30,95	0,00
Ítem 3	0,79	37,89	0,00
Ítem 4	0,80	49,75	0,00
Ítem 5	0,71	52,27	0,00
Ítem 6	0,73	53,76	0,00
Ítem 7	0,78	56,11	0,00
Ítem 8	0,75	53,60	0,00
Ítem 9	0,82	58,72	0,00
Ítem 10	0,85	56,35	0,00
Ítem 11	0,87	51,65	0,00
Ítem 12	0,87	43,31	0,00
Ítem 13	0,90	35,37	0,00

Δ^* Estadístico de contraste

Observamos como en todos los casos rechazamos la hipótesis nula, por lo que $H_i > 0$ para todos los ítems de la Serie B de la Prueba Analogías.

Tabla 14 Índices de Homogeneidad de cada par de ítems (Hij) de la Serie C de la Prueba Analogías

	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9	Ítem 10	Ítem 11	Ítem 12	Ítem 13
Ítem 1	0,89	0,92	0,96	1,00	0,97	1,00	0,92	1,00	0,94	1,00	0,87	0,84
Ítem 2		0,77	0,89	0,96	0,96	1,00	0,95	1,00	0,96	1,00	0,91	0,89
Ítem 3			0,62	0,85	0,90	0,94	0,95	1,00	0,98	1,00	0,95	0,94
Ítem 4				0,67	0,80	0,90	0,95	1,00	0,99	1,00	0,97	0,96
Ítem 5					0,73	0,90	0,96	0,98	0,98	1,00	0,98	0,98
Ítem 6						0,57	0,74	0,85	0,91	0,86	0,85	0,83
Ítem 7							0,70	0,82	0,88	0,92	0,98	0,98
Ítem 8								0,66	0,81	0,89	0,95	0,96
Ítem 9									0,64	0,71	0,81	0,85
Ítem 10										0,87	0,94	0,95
Ítem 11											0,71	0,77
Ítem 12												0,77

Tabla 15 Índices de Homogeneidad de cada Ítem (Hi) y Significación (p) de la Serie C de la Prueba Analogías

	Hi	Δ^*	p
Ítem 1	0,95	32,78	0,00
Ítem 2	0,92	37,15	0,00
Ítem 3	0,86	41,98	0,00
Ítem 4	0,84	47,01	0,00
Ítem 5	0,87	55,91	0,00
Ítem 6	0,77	54,47	0,00
Ítem 7	0,81	63,40	0,00
Ítem 8	0,81	65,37	0,00
Ítem 9	0,77	62,51	0,00
Ítem 10	0,85	68,09	0,00
Ítem 11	0,83	59,06	0,00
Ítem 12	0,85	54,80	0,00
Ítem 13	0,87	51,11	0,00

Δ^* Estadístico de contraste

Observamos como en todos los casos rechazamos la hipótesis nula, por lo que $H_i > 0$ para todos los ítems de la Serie C de la Prueba Analogías.

2.1.4.3. Estudio de los coeficientes H_i y H_{ij} de la prueba Contextos

Tabla 16 Índices de Homogeneidad de cada par de ítems (H_{ij}) de la Serie A de la Prueba Contextos

	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9	Ítem 10	Ítem 11
Ítem 1	0,60	0,66	0,89	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Ítem 2		0,41	0,68	0,62	0,84	0,82	0,93	1,00	1,00	1,00
Ítem 3			0,25	0,46	0,77	0,79	0,81	0,81	0,82	0,93
Ítem 4				0,45	0,84	0,88	0,90	0,95	0,96	1,00
Ítem 5					0,61	0,36	0,38	0,39	0,57	0,86
Ítem 6						0,91	0,94	0,99	0,97	1,00
Ítem 7							0,81	0,82	0,81	0,96
Ítem 8								0,88	0,92	0,85
Ítem 9									0,84	0,93
Ítem 10										0,55

Tabla 17 Índices de Homogeneidad de cada Ítem (H_i) y Significación (p) de la Serie A de la Prueba Contextos

	H_i	Δ^*	p
Ítem 1	0,83	13,26	0,00
Ítem 2	0,70	14,26	0,00
Ítem 3	0,65	42,61	0,00
Ítem 4	0,66	34,67	0,00
Ítem 5	0,45	35,04	0,00
Ítem 6	0,84	62,62	0,00
Ítem 7	0,75	60,11	0,00
Ítem 8	0,77	61,39	0,00
Ítem 9	0,78	53,86	0,00
Ítem 10	0,80	31,63	0,00
Ítem 11	0,85	20,96	0,00

Δ^* Estadístico de contraste

En la anterior tabla se muestra el valor estimado H_i , para cada ítem de la Serie A de la prueba Contextos. Para cada H_i se han planteado las hipótesis:

Hipótesis Nula: $H_i=0$

Hipótesis Alternativa: $H_i>0$

Para muestras de tamaño suficientemente grande, el estadístico del contraste, Δ^* , se distribuye según una distribución normal de media 0 y desviación típica 1 (Mokken, 1971, pág.163) y consideraremos que H_1 es distinto de cero si $\Delta^* \geq Z_{1-\alpha}$.

Observamos como en todos los casos rechazamos la hipótesis nula, por lo que $H_1 > 0$ para todos los ítems de la Serie A de la Prueba Contextos.

Tabla 18 Índices de Homogeneidad de cada par de ítems (H_{ij}) de la Serie B de la Prueba Contextos

	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9	Ítem 10	Ítem 11
Ítem 1	0,48	0,50	0,65	0,79	0,80	0,84	0,92	0,85	0,91	1,00
Ítem 2		0,62	0,71	0,79	0,74	0,78	0,87	0,89	0,89	1,00
Ítem 3			0,49	0,61	0,73	0,90	0,94	0,95	0,95	1,00
Ítem 4				0,62	0,79	0,87	0,96	0,94	0,97	1,00
Ítem 5					0,59	0,74	0,87	0,90	0,94	1,00
Ítem 6						0,47	0,67	0,80	0,87	0,91
Ítem 7							0,49	0,57	0,62	0,73
Ítem 8								0,79	0,91	0,97
Ítem 9									0,71	0,92
Ítem 10										0,69

Tabla 19 Índices de Homogeneidad de cada Ítem (H_i) y Significación (p) de la Serie B de la Prueba Contextos

	H_i	Δ^*	p
Ítem 1	0,70	32,07	0,00
Ítem 2	0,71	29,49	0,00
Ítem 3	0,70	38,02	0,00
Ítem 4	0,74	43,49	0,00
Ítem 5	0,74	48,88	0,00
Ítem 6	0,68	51,26	0,00
Ítem 7	0,60	47,59	0,00
Ítem 8	0,74	58,37	0,00
Ítem 9	0,76	52,63	0,00
Ítem 10	0,80	47,16	0,00
Ítem 11	0,86	26,08	0,00

Δ^* Estadístico de contraste

Observamos como en todos los casos rechazamos la hipótesis nula, por lo que $H_1 > 0$ para todos los ítems de la Serie B de la Prueba Contextos.

Tabla 20 Índices de Homogeneidad de cada par de ítems (Hij) de la Serie C de la Prueba Contextos

	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9	Ítem 10	Ítem 11
Ítem 1	0,61	0,66	0,74	0,85	1,00	0,95	1,00	1,00	0,85	1,00
Ítem 2		0,59	0,69	0,88	1,00	0,96	1,00	1,00	0,88	1,00
Ítem 3			0,71	0,88	0,95	0,94	1,00	1,00	0,88	1,00
Ítem 4				0,65	0,84	0,88	0,87	0,76	0,72	0,59
Ítem 5					0,61	0,78	0,92	0,98	0,97	0,98
Ítem 6						0,75	0,92	0,95	0,98	1,00
Ítem 7							0,78	0,85	0,95	0,98
Ítem 8								0,76	0,89	0,94
Ítem 9									0,41	0,68
Ítem 10										0,66

Tabla 21 Índices de Homogeneidad de cada Ítem (Hi) y Significación (p) de la Serie C de la Prueba Contextos

	Hi	Δ^*	p
Ítem 1	0,80	31,80	0,00
Ítem 2	0,79	33,98	0,00
Ítem 3	0,79	34,32	0,00
Ítem 4	0,76	30,82	0,00
Ítem 5	0,79	51,01	0,00
Ítem 6	0,81	56,07	0,00
Ítem 7	0,82	57,10	0,00
Ítem 8	0,86	56,35	0,00
Ítem 9	0,73	42,49	0,00
Ítem 10	0,77	45,38	0,00
Ítem 11	0,83	34,94	0,00

Δ^* Estadístico de contraste

Observamos como en todos los casos rechazamos la hipótesis nula, por lo que $H_i > 0$ para todos los ítems de la Serie C de la Prueba Contextos.

2.1.5. Monotonía de dificultad de los ítems

Las estimaciones de los coeficientes de escalabilidad hasta ahora calculadas (H , H_i y H_{ij}) se usan para probar la homogeneidad monótona de los ítems. Esta es una condición necesaria, pero no suficiente (Martínez Arias, M.R. y Rivas Moya, T., 1990). La prueba de monotonía de las dificultades de los ítems supone realizar una inspección de las matrices $P(+,+)$ y $P(-,-)$, que contienen las probabilidades de dos respuestas positivas y dos respuestas negativas simultáneamente, para todos los pares posibles.

A continuación se ofrecen las matrices $P(+,+)$ (que representa la proporción de sujetos que han acertado correctamente cada par de ítems) y $P(-,-)$ (que representa la proporción de sujetos que han fallado cada par de ítems). Ambas matrices son simétricas, por lo que sólo presentamos la mitad triangular superior de cada una de ellas.

Las puntuaciones obtenidas en todos los pares forman una escala creciente o estable, excepto en los valores señalados con un asterisco.

2.1.5.1. Monotonía de las dificultades de los ítems en la prueba Semejanzas

Tabla 22 Matriz $P(+,+)$ de la Serie A de la Prueba Semejanzas

	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9
Ítem 1		0,99	0,99	0,99	0,99	0,99	1,00	1,00	1,00
Ítem 2			0,97	0,98	0,99	1,00	1,00	1,00	1,00
Ítem 3				0,66	0,75	0,78	0,83	0,89	1,00
Ítem 4					0,73	0,77	0,79	0,84	0,92
Ítem 5						0,88	0,91	0,95	1,00
Ítem 6							0,86	0,90	1,00
Ítem 7								0,78	1,00
Ítem 8									0,38
Ítem 9									
ID	0,98	0,95	0,64	0,60	0,60	0,47	0,22	0,04	0,01

Tabla 23 Matriz P (+, +) de la Serie B de la Prueba Semejanzas

	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9
Ítem 1		0,99	0,99	0,99	1,00*	1,00*	0,99	0,99	0,99
Ítem 2			0,99	0,99	1,00*	1,00*	0,99	1,00	1,00
Ítem 3				0,90	0,95	0,97	0,97	0,98	0,98
Ítem 4					0,91	0,94	0,95	0,98	0,99
Ítem 5						0,73	0,83	0,88	0,94
Ítem 6							0,80	0,88	0,92
Ítem 7								0,79	0,87
Ítem 8									0,73
Ítem 9									
ID	0,97	0,97	0,83	0,78	0,49	0,44	0,31	0,16	0,08

Tabla 24 Matriz P (+, +) de la Serie C de la Prueba Semejanzas

	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9
Ítem 1		0,99	0,99	1,00	1,00	1,00	1,00	1,00	1,00
Ítem 2			0,97	0,98	0,99	0,99	0,99	0,99	1,00
Ítem 3				0,96	0,98	0,99	0,99	1,00	1,00
Ítem 4					0,92	0,95	0,98	0,98	0,99
Ítem 5						0,83	0,90	0,95	1,00
Ítem 6							0,84	0,90	0,95
Ítem 7								0,71	0,87
Ítem 8									0,68
Ítem 9									
ID	0,97	0,93	0,90	0,81	0,64	0,57	0,35	0,19	0,07

Tabla 25 Matriz P (-,-) de la Serie A de la Prueba Semejanzas

	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9
Ítem 1		0,51	0,67	0,77	0,82	0,91	0,98	1,00	1,00
Ítem 2			0,54	0,74	0,91	0,96	0,98	1,00	1,00
Ítem 3				0,44	0,58	0,71	0,89	0,99	1,00
Ítem 4					0,59	0,73	0,88	0,98	1,00
Ítem 5						0,85	0,95	0,99	1,00
Ítem 6							0,94	0,99	1,00
Ítem 7								0,99	1,00
Ítem 8									0,99
Ítem 9									
ID	0,98	0,95	0,64	0,60	0,60	0,47	0,22	0,04	0,01

Tabla 26 Matriz P (-,-) de la Serie B de la Prueba Semejanzas

	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9
Ítem 1		0,57	0,72	0,77	0,93	0,95	0,93	0,97	0,98
Ítem 2			0,76	0,77	0,93	0,96	0,95	0,99	1,00
Ítem 3				0,56	0,85	0,92	0,95	0,98	0,99
Ítem 4					0,80	0,88	0,93	0,99	1,00
Ítem 5						0,77	0,90	0,96	0,99
Ítem 6							0,89	0,96	0,99
Ítem 7								0,95	0,98
Ítem 8									0,97
Ítem 9									
ID	0,97	0,97	0,83	0,78	0,49	0,44	0,31	0,16	0,08

Tabla 27 Matriz P (-,-) de la Serie C de la Prueba Semejanzas

	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9
Ítem 1		0,74	0,77	0,89	0,94	0,98	0,97	0,98	1,00
Ítem 2			0,58	0,77	0,90	0,94	0,97	0,98	1,00
Ítem 3				0,69	0,85	0,93	0,97	1,00	1,00
Ítem 4					0,74	0,85	0,96	0,98	1,00
Ítem 5						0,73	0,90	0,97	1,00
Ítem 6							0,87	0,96	0,99
Ítem 7								0,91	0,99
Ítem 8									0,97
Ítem 9									
ID	0,97	0,93	0,90	0,81	0,64	0,57	0,35	0,19	0,07

2.1.5.2. Monotonía de las dificultades de los ítems en la prueba Analogías

Tabla 28 Matriz P (+, +) de la Serie A de la Prueba Analogías

	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9	Ítem 10	Ítem 11	Ítem 12	Ítem 13
Ítem 1		0,99	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Ítem 2			0,64	0,67	0,73	0,79	0,86	0,90	0,92	0,92	0,99	0,99	0,98*
Ítem 3				0,96	0,97	0,98	0,99	0,99	1,00	1,00	0,99*	1,00	1,00
Ítem 4					0,93	0,94	0,97	0,99	0,99	0,99	0,99	0,99	1,00
Ítem 5						0,90	0,93	0,95	0,97	0,98	0,99	0,99	1,00
Ítem 6							0,79	0,84	0,89	0,89	0,92	0,95	0,98
Ítem 7								0,83	0,84	0,89	0,92	0,92	0,95
Ítem 8									0,85	0,87	0,93	0,94	0,91*
Ítem 9										0,89	0,92	0,94	0,98
Ítem 10											0,70	0,78	0,86
Ítem 11												0,81	0,81
Ítem 12													0,86
Ítem 13													
ID	0,97	0,60	0,88	0,77	0,64	0,45	0,34	0,25	0,22	0,12	0,07	0,05	0,02

Tabla 29 Matriz P (+, +) de la Serie B de la Prueba Analogías

	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9	Ítem 10	Ítem 11	Ítem 12	Ítem 13
Ítem 1		0,99	0,99	1,00*	1,00*	0,99	0,99	0,99	0,99	0,99	1,00	0,99	1,00
Ítem 2			0,99	0,99	0,99	0,99	0,98*	0,99	0,99	0,99	1,00*	0,99	1,00
Ítem 3				0,97	0,98	0,98	0,98	0,99	0,99	0,99	1,00*	0,99	1,00
Ítem 4					0,93	0,95	0,97	0,99	0,99	0,99	1,00*	0,99	1,00
Ítem 5						0,83	0,89	0,95	0,97	0,99	0,99	0,99	1,00
Ítem 6							0,89	0,95	0,98	0,99	1,00*	0,99	1,00
Ítem 7								0,87	0,94	0,96	0,97	0,98	1,00
Ítem 8									0,61	0,77	0,84	0,89	0,92
Ítem 9										0,86	0,90	0,93	0,99
Ítem 10											0,83	0,92	0,93
Ítem 11												0,78	0,84
Ítem 12													0,75
Ítem 13													
ID	0,96	0,96	0,92	0,81	0,68	0,66	0,50	0,23	0,25	0,15	0,10	0,06	0,03

Tabla 30 Matriz P (+, +) de la Serie C de la Prueba Analogías

	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9	Ítem 10	Ítem 11	Ítem 12	Ítem 13
Ítem 1		1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Ítem 2			0,99	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Ítem 3				0,97	0,99	0,99	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Ítem 4					0,96	0,98	0,99	0,99	1,00	1,00	1,00	1,00	1,00
Ítem 5						0,95	0,98	0,99	1,00	1,00	1,00	1,00	1,00
Ítem 6							0,85	0,91	0,95	0,97	0,98	0,98	0,98
Ítem 7								0,87	0,92	0,95	0,97	0,99	0,99
Ítem 8									0,81	0,90	0,94	0,97	0,98
Ítem 9										0,75	0,80	0,87	0,90
Ítem 10											0,91	0,96	0,97
Ítem 11												0,77	0,81
Ítem 12													0,80
Ítem 13													
ID	0,97	0,96	0,93	0,89	0,81	0,66	0,57	0,44	0,31	0,30	0,19	0,13	0,11

Tabla 31 Matriz P (-,-) de la Serie A de la Prueba Analogías

	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9	Ítem 10	Ítem 11	Ítem 12	Ítem 13
Ítem 1		0,83	0,88	0,93	0,97	0,97	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Ítem 2			0,21	0,37	0,57	0,77	0,88	0,94	0,96	0,98	1,00	1,00	1,00
Ítem 3				0,76	0,85	0,92	0,97	0,99	0,99	1,00	1,00	1,00	1,00
Ítem 4					0,79	0,89	0,95	0,99	0,99	0,99	1,00	1,00	1,00
Ítem 5						0,87	0,94	0,97	0,98	0,99	1,00	1,00	1,00
Ítem 6							0,87	0,93	0,96	0,98	0,99	1,00	1,00
Ítem 7								0,93	0,95	0,98	0,99	0,99	1,00
Ítem 8									0,96	0,98	0,99	1,00	1,00
Ítem 9										0,98	0,99	1,00	1,00
Ítem 10											0,98	0,99	1,00
Ítem 11												0,99	1,00
Ítem 12													1,00
Ítem 13													
ID	0,97	0,60	0,88	0,77	0,64	0,45	0,34	0,25	0,22	0,12	0,07	0,05	0,02

Tabla 32 Matriz P (-,-) de la Serie B de la Prueba Analogías

	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9	Ítem 10	Ítem 11	Ítem 12	Ítem 13
Ítem 1		0,80	0,85	0,93	0,95	0,81	0,82	0,94	0,93	0,98	1,00	0,99	1,00
Ítem 2			0,80	0,88	0,90	0,76	0,81	0,94	0,93	0,98	1,00	0,99	1,00
Ítem 3				0,72	0,84	0,85	0,89	0,97	0,97	0,99	1,00	0,99	1,00
Ítem 4					0,76	0,84	0,92	0,99	0,98	1,00	1,00	1,00	1,00
Ítem 5						0,66	0,82	0,97	0,98	0,99	1,00	1,00	1,00
Ítem 6							0,84	0,96	0,99	1,00	1,00	1,00	1,00
Ítem 7								0,94	0,97	0,99	0,99	1,00	1,00
Ítem 8									0,87	0,95	0,98	0,99	1,00
Ítem 9										0,97	0,99	0,99	1,00
Ítem 10											0,98	0,99	1,00
Ítem 11												0,99	0,99
Ítem 12													0,99
Ítem 13													
ID	0,96	0,96	0,92	0,81	0,68	0,66	0,50	0,23	0,25	0,15	0,10	0,06	0,03

Tabla 33 Matriz P (-,-) de la Serie C de la Prueba Analogías

	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9	Ítem 10	Ítem 11	Ítem 12	Ítem 13
Ítem 1		0,89	0,93	0,96	1,00	0,98	1,00	0,96	1,00	0,98	1,00	0,98	0,98
Ítem 2			0,79	0,90	0,96	0,98	1,00	0,98	1,00	0,99	1,00	0,99	0,99
Ítem 3				0,66	0,88	0,93	0,97	0,98	1,00	0,99	1,00	0,99	0,99
Ítem 4					0,73	0,87	0,94	0,98	1,00	1,00	1,00	1,00	1,00
Ítem 5						0,82	0,94	0,98	1,00	1,00	1,00	1,00	1,00
Ítem 6							0,75	0,88	0,95	0,97	0,99	0,99	0,99
Ítem 7								0,87	0,95	0,97	0,98	1,00	1,00
Ítem 8									0,90	0,95	0,98	0,99	1,00
Ítem 9										0,89	0,95	0,97	0,98
Ítem 10											0,98	0,99	0,99
Ítem 11												0,96	0,98
Ítem 12													0,97
Ítem 13													
ID	0,97	0,96	0,93	0,89	0,81	0,66	0,57	0,44	0,31	0,30	0,19	0,13	0,11

2.1.5.3. Monotonía de las dificultades de los ítems en la prueba Contextos

Tabla 34 Matriz P (+, +) de la Serie A de la Prueba Contextos

	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9	Ítem 10	Ítem 11
Ítem 1		1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Ítem 2			0,99	0,99	0,99	1,00	1,00	1,00	1,00	1,00	1,00
Ítem 3				0,69	0,81	0,92	0,92	0,93	0,93	0,93	0,97
Ítem 4					0,88	0,96	0,97	0,98	0,99	0,99	1,00
Ítem 5						0,43	0,45	0,52	0,55	0,69	0,89
Ítem 6							0,96	0,97	0,99	0,98	1,00
Ítem 7								0,88	0,88	0,88	0,97
Ítem 8									0,91	0,94	0,89
Ítem 9										0,88	0,95
Ítem 10											0,58
Ítem 11											
ID	0,99	0,98	0,64	0,78	0,27	0,51	0,36	0,30	0,20	0,06	0,02

Tabla 35 Matriz P (+, +) de la Serie B de la Prueba Contextos

	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9	Ítem 10	Ítem 11
Ítem 1		0,97	0,97	0,98	0,99	0,99	0,99	1,00*	0,99	1,00	1,00
Ítem 2			0,98	0,99	0,99	0,99	0,99	0,99	1,00	1,00	1,00
Ítem 3				0,95	0,96	0,98	0,99	0,99	1,00	1,00	1,00
Ítem 4					0,95	0,97	0,98	0,99	0,99	1,00	1,00
Ítem 5						0,90	0,94	0,97	0,98	0,98	1,00
Ítem 6							0,77	0,86	0,91	0,95	0,96
Ítem 7								0,73	0,77	0,80	0,85
Ítem 8									0,88	0,95	0,98
Ítem 9										0,80	0,94
Ítem 10											0,75
Ítem 11											
ID	0,95	0,96	0,91	0,87	0,76	0,57	0,46	0,44	0,30	0,21	0,05

Tabla 36 Matriz P (+, +) de la Serie C de la Prueba Contextos

	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9	Ítem 10	Ítem 11
Ítem 1		0,99	0,99	0,99	0,99	1,00	1,00	1,00	1,00	0,99	1,00
Ítem 2			0,98	0,99	0,99	1,00	1,00	1,00	1,00	0,99	1,00
Ítem 3				0,99	0,99	1,00	1,00	1,00	1,00	0,99	1,00
Ítem 4					0,98	0,99	1,00	1,00	1,00	0,99	1,00
Ítem 5						0,87	0,93	0,97	0,99	0,99	0,99
Ítem 6							0,90	0,97	0,98	0,99	1,00
Ítem 7								0,89	0,93	0,98	0,99
Ítem 8									0,84	0,93	0,96
Ítem 9										0,48	0,73
Ítem 10											0,71
Ítem 11											
ID	0,96	0,95	0,95	0,92	0,68	0,60	0,50	0,32	0,16	0,17	0,07

Tabla 37 Matriz P (-,-) de la Serie A de la Prueba Contextos

	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9	Ítem 10	Ítem 11
Ítem 1		0,61	0,78	0,91	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Ítem 2			0,63	0,75	0,90	0,92	0,94	0,98	1,00	1,00	1,00
Ítem 3				0,32	0,85	0,88	0,93	0,94	0,96	0,99	1,00
Ítem 4					0,85	0,92	0,96	0,97	0,99	1,00	1,00
Ítem 5						0,60	0,73	0,80	0,88	0,98	1,00
Ítem 6							0,97	0,98	1,00	1,00	1,00
Ítem 7								0,94	0,96	0,99	1,00
Ítem 8									0,98	1,00	1,00
Ítem 9										0,99	1,00
Ítem 10											0,99
Ítem 11											
ID	0,99	0,98	0,64	0,78	0,27	0,51	0,36	0,30	0,20	0,06	0,02

Tabla 38 Matriz P (-,-) de la Serie B de la Prueba Contextos

	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9	Ítem 10	Ítem 11
Ítem 1		0,40	0,55	0,69	0,84	0,88	0,93	0,96	0,95	0,98	1,00
Ítem 2			0,66	0,75	0,84	0,85	0,90	0,94	0,97	0,98	1,00
Ítem 3				0,56	0,70	0,85	0,96	0,98	0,99	0,99	1,00
Ítem 4					0,72	0,88	0,94	0,98	0,98	0,99	1,00
Ítem 5						0,77	0,88	0,94	0,97	0,99	1,00
Ítem 6							0,76	0,86	0,94	0,97	1,00
Ítem 7								0,78	0,87	0,92	0,99
Ítem 8									0,94	0,98	1,00
Ítem 9										0,94	1,00
Ítem 10											0,98
Ítem 11											
ID	0,95	0,96	0,91	0,87	0,76	0,57	0,46	0,44	0,30	0,21	0,05

Tabla 39 Matriz P (-,-) de la Serie C de la Prueba Contextos

	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9	Ítem 10	Ítem 11
Ítem 1		0,63	0,68	0,76	0,90	1,00	0,98	1,00	1,00	0,98	1,00
Ítem 2			0,61	0,72	0,92	1,00	0,98	1,00	1,00	0,98	1,00
Ítem 3				0,73	0,92	0,97	0,97	1,00	1,00	0,98	1,00
Ítem 4					0,81	0,95	0,98	1,00	1,00	0,99	1,00
Ítem 5						0,77	0,89	0,97	1,00	0,99	1,00
Ítem 6							0,87	0,97	0,99	1,00	1,00
Ítem 7								0,93	0,98	0,99	1,00
Ítem 8									0,96	0,98	1,00
Ítem 9										0,90	0,98
Ítem 10											0,97
Ítem 11											
ID	0,96	0,95	0,95	0,92	0,68	0,60	0,50	0,32	0,16	0,17	0,07

2.2. Distribuciones de puntuaciones

La **Batería CERVANTES** consta de tres pruebas generales a partir de cuyos resultados se obtiene el intervalo correspondiente al CI o a la puntuación centil: Semejanzas, Analogías y Contextos. La prueba de Diseños se propone como prueba complementaria de 6 a 12 años. La prueba Semejanzas consta de 27 ítems, por lo que la puntuación obtenida en esta prueba varía de 0 a 27. La prueba Analogías consta de 39 ítems y la prueba contextos consta de 33 ítems. La puntuación global de un sujeto varía entre una puntuación mínima de 0 a una máxima de 99. La prueba Diseños consta de 18 ítems.

Tabla 40 Correlaciones entre pruebas por edad

Edad	Correlación entre pruebas
6 Años	0,48
7 Años	0,52
8 Años	0,54
9 Años	0,48
10 Años	0,57
11 Años	0,44
12 Años	0,60
13 Años	0,49
14 Años	0,51
15 Años	0,48
16 Años	0,45
17 Años	0,41
18 Años	0,63

Tabla 41 Puntuación media por edad en cada prueba

	Media		
	Semejanzas	Analogías	Contextos
6 Años	8,99	9,45	9,32
7 Años	11,61	11,83	11,89
8 Años	12,92	14,68	13,46
9 Años	13,94	16,33	15,76
10 Años	14,64	17,80	17,34
11 Años	15,60	18,68	18,38
12 Años	15,96	20,11	20,02
13 Años	15,77	20,72	19,52
14 Años	16,31	22,38	20,61
15 Años	17,66	24,08	22,19
16 Años	17,98	25,79	22,72
17 Años	18,78	27,74	23,67
18 Años	18,52	26,64	24,52

Tabla 42 Mediana y rango Intercuartílico por edad para cada prueba

	Mediana			Rango Intercuartílico*		
	Semejanzas	Analogías	Contextos	Semejanzas	Analogías	Contextos
6 Años	9	9.5	9	6 - 12	6 - 12	6.75 - 12
7 Años	12	12	12	9 - 15	9 - 15	10 - 14
8 Años	13	14	13	10 - 16	11 - 18	10 - 16
9 Años	14	16	15	11 - 17	13 - 20	13 - 19
10 Años	15	17	17	12 - 17	13 - 21.25	13 - 21
11 Años	16	19	18	13 - 19	15 - 22	14 - 22
12 Años	16	20	19	13 - 19	16 - 24	15 - 24.5
13 Años	16	21	19	13 - 19	16 - 25	15 - 25
14 Años	16	21	21	13.5 - 19	18 - 28	16 - 25
15 Años	18	23	23	15 - 21	20 - 29	18.25 - 26
16 Años	18	25	24	15 - 21	21 - 31	20 - 27
17 Años	20	27	24	16 - 22	22 - 33.75	21 - 27
18 Años	20	28	24	15 - 21	21 - 31	22 - 28

*Distancia entre el percentil 25 y percentil 75

La mediana de las puntuaciones en las tres pruebas tiene un incremento monótono con la edad, mayor en los grupos de 6 a 12 años, de 7, 10 y 11 puntos en las tres pruebas y es de 2, 7 y 5 puntos en las tres pruebas de 12 a 18 años.

La desviación estándar no tiene relación lineal con la edad.

El 50% de los casos centrales (rango intercuartílico) se sitúan en un rango de ± 3 aciertos, intervalo que no es constante en todos los grupos de edad.

Figura 1 Distribución de las puntuaciones directas de la Prueba Semejanzas de 6 a 8 años

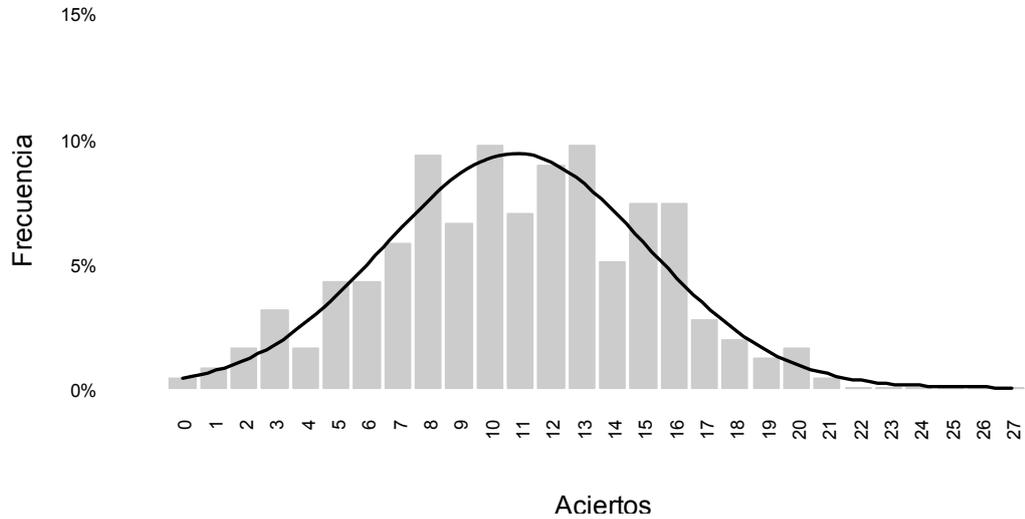


Figura 2 – Distribución de las puntuaciones directas de la Prueba Semejanzas de 8 a 11 años

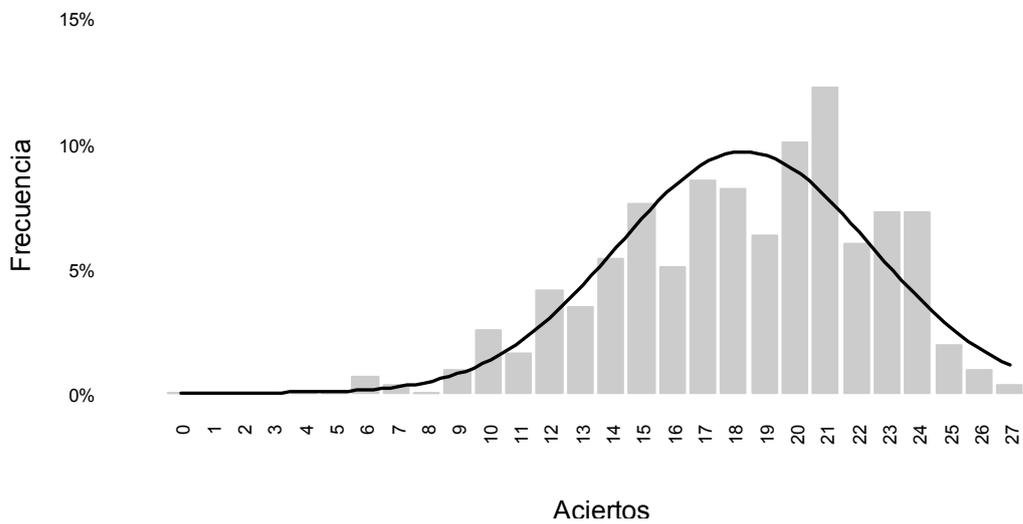


Figura 3 – Distribución de las puntuaciones directas de la Prueba Semejanzas de 11 a 15 años

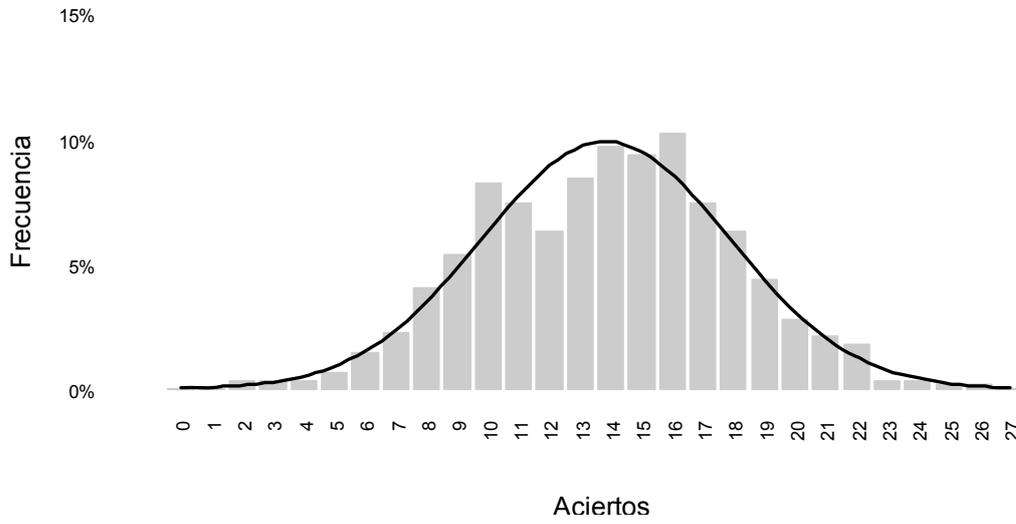


Figura 4 – Distribución de las puntuaciones directas de la Prueba Semejanzas para mayores de 15 años

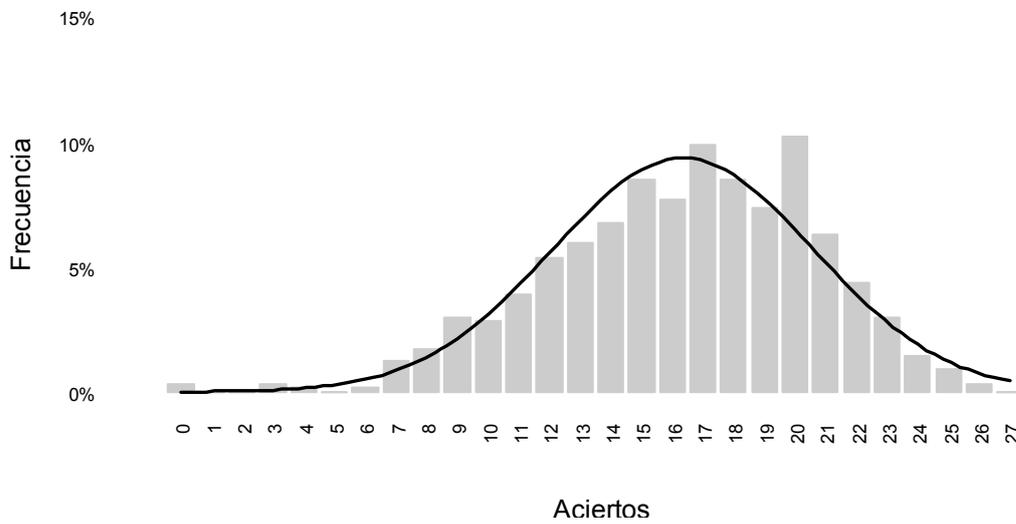


Figura 5 – Distribución de las puntuaciones directas de la Prueba Analogías de 6 a 8 años

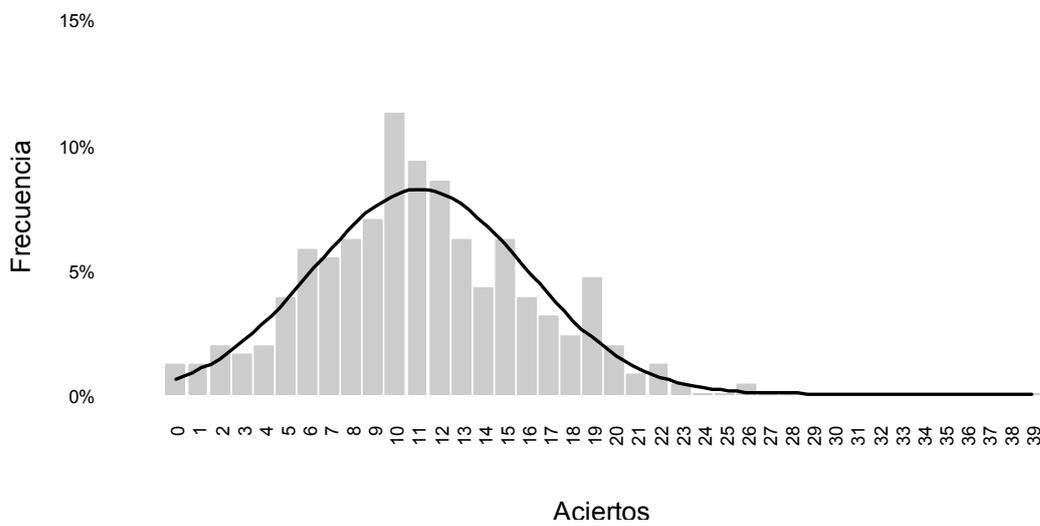


Figura 6 – Distribución de las puntuaciones directas de la Prueba Analogías de 8 a 11 años

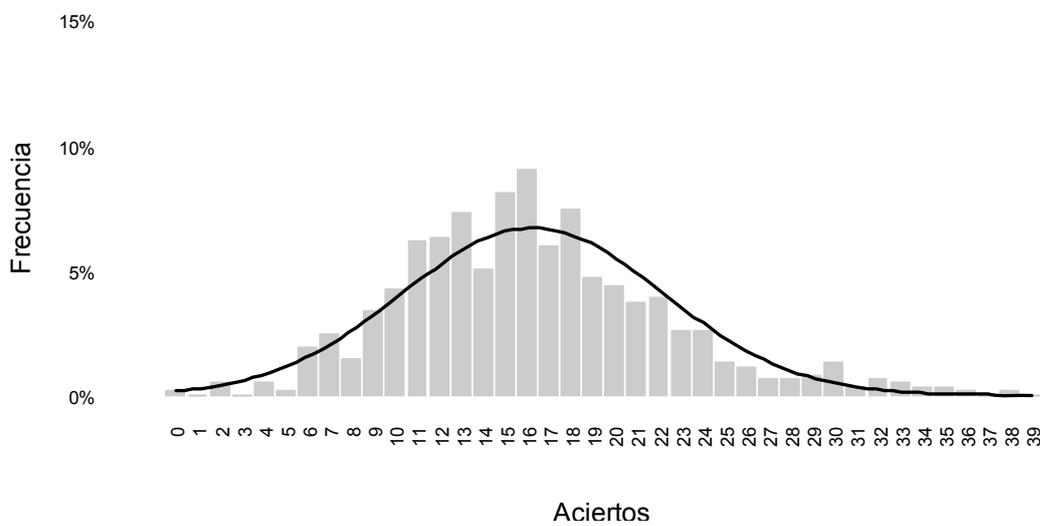


Figura 7 – Distribución de las puntuaciones directas de la Prueba Analogías de 11 a 15 años

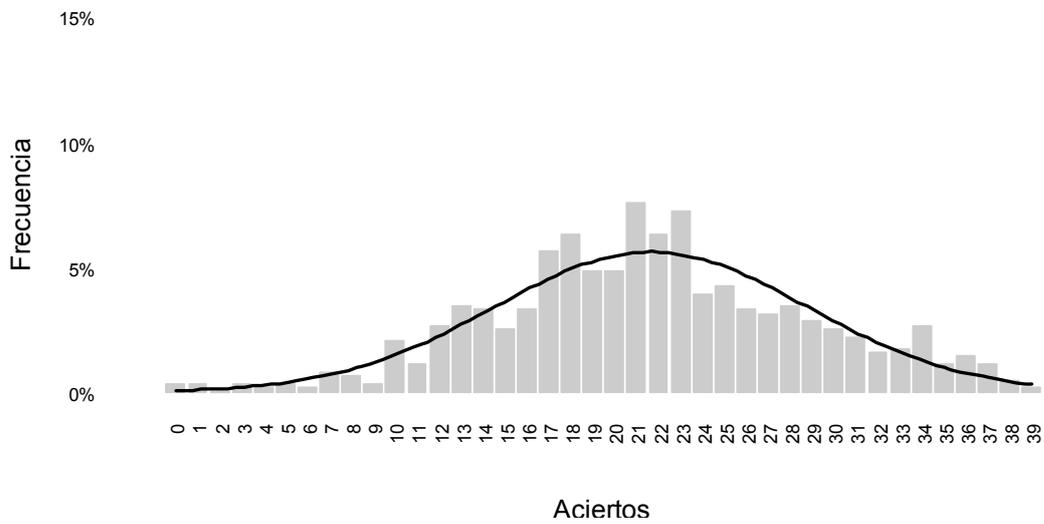


Figura 8 – Distribución de las puntuaciones directas de la Prueba Analogías para mayores de 15 años

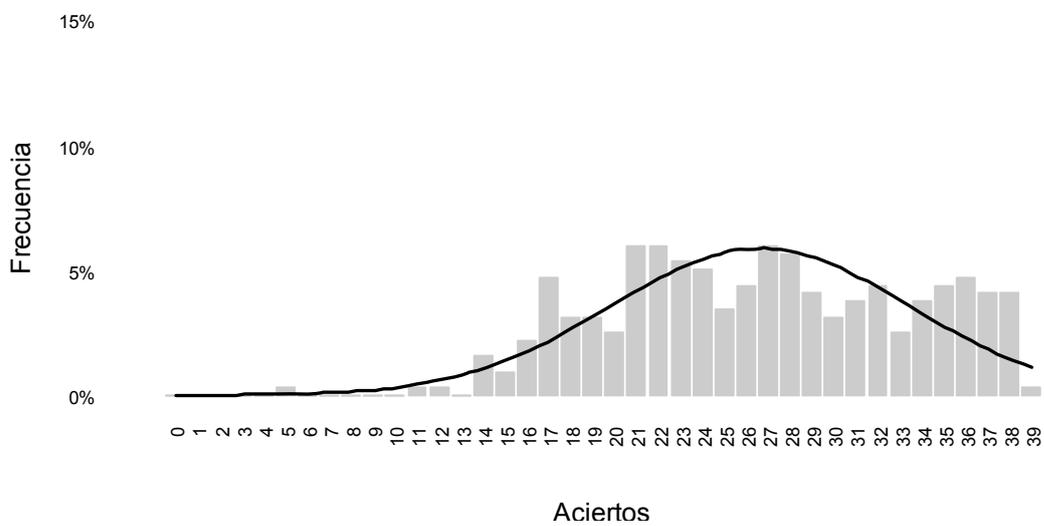


Figura 9 – Distribución de las puntuaciones directas de la Prueba Contextos de 6 a 8 años

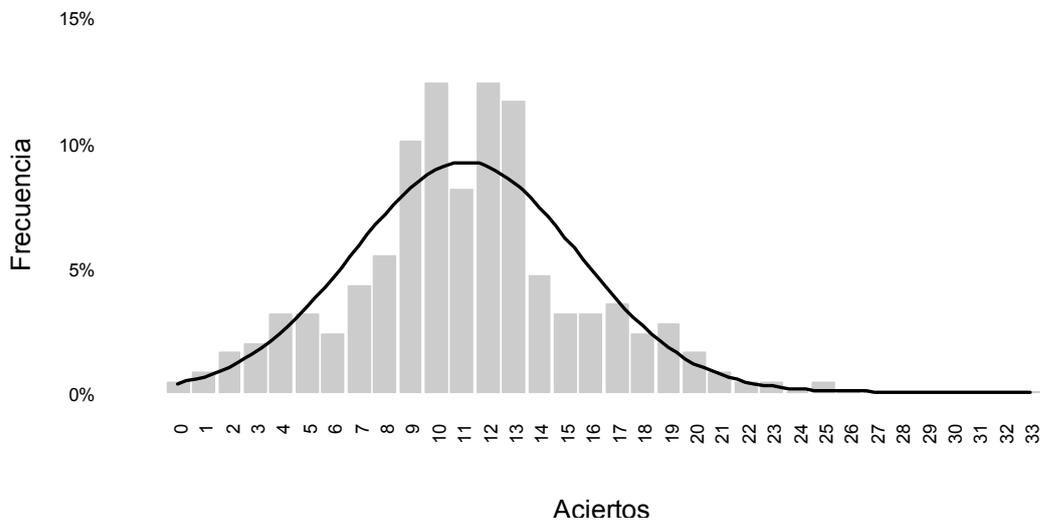


Figura 10 – Distribución de las puntuaciones directas de la Prueba Contextos de 8 a 11 años

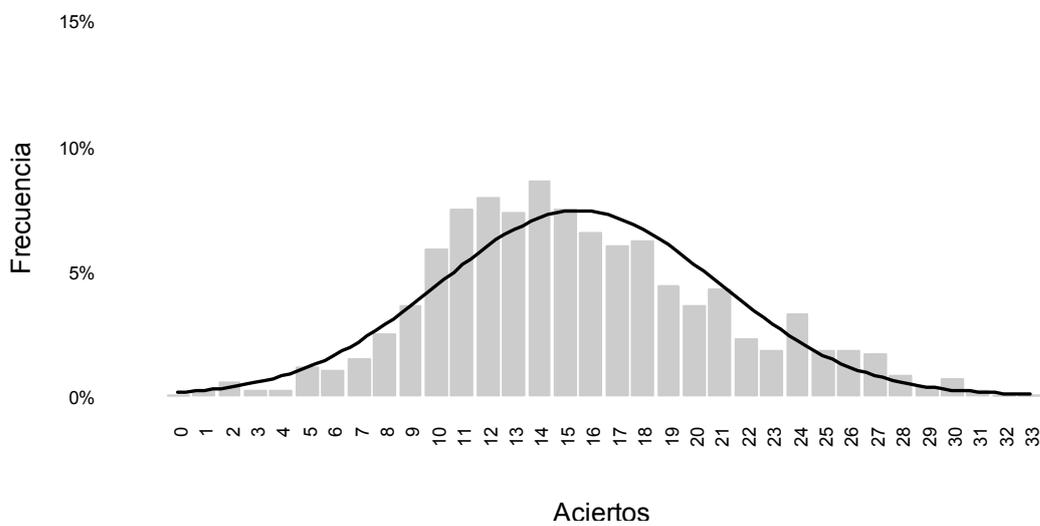


Figura 11 – Distribución de las puntuaciones directas de la Prueba Contextos de 11 a 15 años

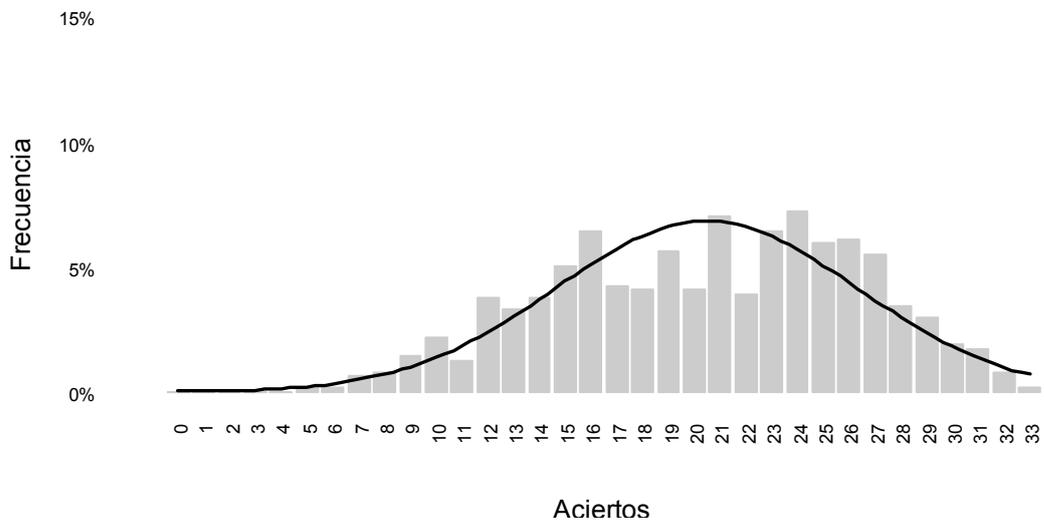
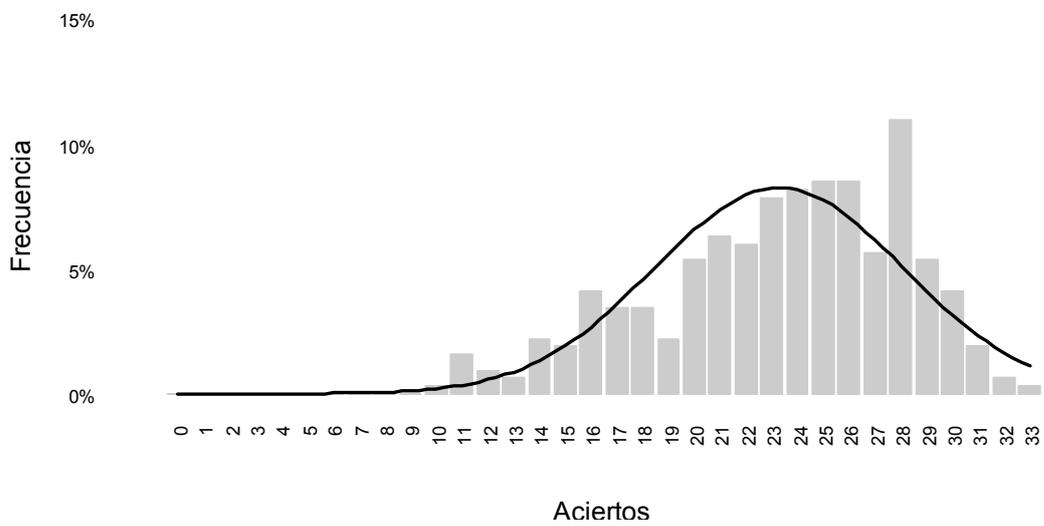


Figura 12 – Distribución de las puntuaciones directas de la Prueba Contextos para mayores de 15 años



3. Muestras de Baremación

3.1. Distribución por edades

La muestra definitiva de la baremación incluye 1.983 sujetos que se han distribuido en 18 grupos por edad que van desde los 6 a los 18 años de 6 en 6 meses.

Tabla 43 Distribución de la población por edad y sexo

Años cumplidos	Femenino	Masculino	Total	% Femenino	% Masculino
6	35	39	74	3,5%	4,0%
7	90	92	182	8,9%	9,4%
8	98	101	199	9,7%	10,4%
9	111	107	218	11,0%	11,0%
10	96	103	199	9,5%	10,6%
11	102	76	178	10,1%	7,8%
12	80	88	168	7,9%	9,0%
13	85	91	176	8,4%	9,3%
14	80	70	150	7,9%	7,2%
15	63	69	132	6,3%	7,1%
16	86	64	150	8,5%	6,6%
17	65	59	124	6,4%	6,1%
18	17	16	33	1,7%	1,6%
Total	1.008	975	1.983		

3.2. Distribución geográfica

Se han recogido casos de 12 Comunidades Autónomas.

Tabla 44 Distribución de la población por Comunidad Autónoma de residencia

Comunidad	Total	%
Andalucía	336	16,9%
Aragón	89	4,5%
Baleares Islas	36	1,8%
Canarias	30	1,5%
Cantabria	45	2,3%
Castilla y León	20	1,0%
Castilla-La Mancha	106	5,3%
Comunidad Valenciana	158	8,0%
Galicia	366	18,5%
Madrid (Comunidad de)	212	10,7%
Murcia (Región de)	147	7,4%
País Vasco	438	22,1%
	1.983	

3.3. Centros educativos

Se han recogido casos de 37 Centros de Educación Primaria.

Tabla 45 Distribución de los Centros de Educación Primaria

Centro	Edad	Educación Primaria						
		6	7	8	9	10	11	12
Academia Plaza – Centro de Formación		0	0	5	3	3	0	4
APAE		8	11	11	11	11	9	9
Centro Altatorre		0	5	5	2	0	5	3
Luis de Eleizalde Ikastola		0	11	9	7	6	9	5
C. ALAZNE		0	1	5	4	2	4	0
C.P. Fernando Martín		4	7	4	7	5	6	3
CEIP Isaac Peral		1	9	2	10	12	9	2
CEIP Las Esperanzas		0	9	0	0	6	14	4
CEIP Luis Cernuda		3	8	5	5	6	2	5
CEIP Pons Sorolla		5	4	3	3	1	2	3
CEIP Xuño		1	9	5	6	8	2	4
Centro Psicopedagógico Educas		5	1	4	2	2	0	6
Colegio Amor de Dios		2	5	7	5	4	5	6
Colegio Angel de la Guarda		0	0	5	2	9	8	1
Colegio CEU San Pablo Montepíncipe		8	1	5	10	9	2	0
Colegio Espíritu Santo		3	6	4	7	3	5	6
Colegio Hijas de Cristo Rey		1	4	3	4	6	6	4
Colegio La Higuera		2	3	5	7	6	4	3
Colegio La Inmaculada		2	5	7	8	5	7	1
Colegio la Milagrosa - Josefa Sobrido		3	6	8	3	6	7	6
Colegio Nazaret		0	3	4	2	4	4	4
Colegio Ntra. Sra. del Rosario		4	10	9	9	3	0	0
Colegio Ntra. Sra. Gracia		3	5	8	4	8	3	5
Colegio Obradoiro		0	6	3	5	5	5	5
Colegio Público El Greco		2	6	7	6	3	8	2
Colegio San Alberto		0	6	7	6	6	6	9
Colegio San Rafael		0	0	0	4	5	4	5
Colegio Santa Ana		2	5	6	4	6	7	1
Colegio Villa de Griñón		1	2	3	8	3	2	2
Colegio Virgen del Carmen		0	0	6	14	9	3	0
Colegio Vizcaya		4	11	6	9	6	9	7
CP Amara-Berri		0	0	12	17	7	0	0
El Limonar Internacional School		2	5	5	6	6	4	2
Niño Jesús de Praga Karmelo Ikastetxea		2	6	7	2	6	5	4
San Bizente Ikastola		3	4	5	6	4	6	5
Santísima Trinidad I y II		4	3	7	5	3	6	1
Colegio Urkide		1	6	4	5	6	2	3

Se han recogido casos de 37 centros de Educación Secundaria.

Tabla 46 Distribución de los Centros de Educación Secundaria

Centro	Edad	Educación Secundaria					
		13	14	15	16	17	18
Academia Plaza – Centro de Formación		4	2	0	1	0	0
APAE		5	8	3	3	2	0
Askartza-Claret		14	15	7	4	0	0
Luís de Eleizalde Ikastola		10	5	7	5	0	0
CEIP Las Esperanzas		1	0	0	0	0	0
CEIP Xuño		1	0	0	0	0	0
Centro Psicopedagógico Educas		9	3	1	0	1	1
Colegio Amor de Dios		1	0	0	0	0	0
Colegio Angel de la Guarda		8	6	3	4	0	0
Colegio Espíritu Santo		1	0	0	0	0	0
Colegio Hijas de Cristo Rey		3	1	2	1	0	0
Colegio la Milagrosa - Josefa Sobrido		6	5	7	2	0	0
Colegio Montfort		0	0	0	2	17	5
Colegio Nazaret		5	6	3	3	0	0
Colegio Obradoiro		4	6	0	0	0	0
Colegio San Alberto		3	6	3	4	0	0
Colegio San Rafael		7	4	3	3	0	0
Colegio Santa Ana		4	6	7	5	1	1
Colegio Villa de Griñón		0	2	1	1	17	0
Colegio Virgen del Carmen		6	6	3	7	14	2
Colegio Vizcaya		8	10	7	7	5	3
Cristina Bravo		0	0	2	28	24	3
Divina Pastora - Calasancios		12	6	12	3	2	0
Erandio BHI		15	7	10	4	0	0
Escuela San José		11	3	9	4	4	4
IES Félix de Azara		12	5	10	6	2	0
IES Foramontanos		17	10	10	6	2	0
IES Humanejos		8	1	5	17	9	4
IES Los Molinos		15	8	4	1	0	0
IES Mario López		0	0	2	12	2	0
IES Ramón y Cajal		6	7	5	6	6	5
IES Terra de Trasancos		4	4	0	2	0	0
Colegio María Inmaculada		0	0	0	5	11	1
Niño Jesús de Praga Karmelo Ikastetxea		7	6	1	4	0	0
San Bizente Ikastola		5	2	5	4	0	0
Santísima Trinidad I y II		7	5	2	6	4	2
Colegio Urkide		0	0	0	0	3	2

4. Estandarización de la distribución de las puntuaciones

Para la baremación de la **Batería CERVANTES**, se ha considerado la distribución seguida por las puntuaciones, por edad, para obtener un ajuste de las curvas seguidas por las puntuaciones directas obtenidas en cada prueba para cada grupo de edad. Para ello hemos tipificado las puntuaciones siguiendo el modelo:

$$\begin{aligned}
 Z = & A_1 \cdot PD_{SEMEJANZAS} \cdot Edad + \\
 & A_2 \cdot PD_{ANALOGÍAS} \cdot Edad + \\
 & A_3 \cdot PD_{CONTEXTOS} \cdot Edad + \\
 & A_4 \cdot PD_{SEMEJANZAS} \cdot PD_{ANALOGÍAS} \cdot PD_{CONTEXTOS}
 \end{aligned}$$

Donde Z es el valor tipificado de la estimación de una puntuación directa y edad en el momento de la aplicación. Los parámetros A_1 , A_2 , A_3 y A_4 se han determinado de tal manera que minimizan las diferencias entre los valores reales obtenidos con la muestra y los valores obtenidos con el modelo.

Hemos optado por este modelo ya que considera las puntuaciones directas obtenidas en cada prueba según la edad en el momento de la aplicación (los tres primeros términos del modelo) y además considera la relación existente entre las puntuaciones directas obtenidas en las tres pruebas (último término del modelo)

El término del modelo que considera la relación entre las tres puntuaciones es importante ya que nos permite discriminar entre aquellos sujetos que han tenido una misma puntuación directa como suma de las tres pruebas (0-99) pero distintas puntuaciones directas en cada prueba (Semejanzas 0-27, Analogías 0-39, Contextos 0-33). Esto nos ha permitido, además, considerar el concepto de "*perfil homogéneo*" y perfil "*heterogéneo de resultados*", para reducir los posibles sesgos de factores que comprometan la fiabilidad de la prueba.

Los valores Z son calculados en base a las proporciones acumuladas de las distribuciones discretas de las puntuaciones directas obtenidas en cada prueba, según la edad en el momento de la aplicación. El valor tipificado final se construye en base a las proporciones acumuladas de la distribución discreta de la suma de las puntuaciones directas.

Para cada puntuación directa obtenida, los valores obtenidos Z cumplen:

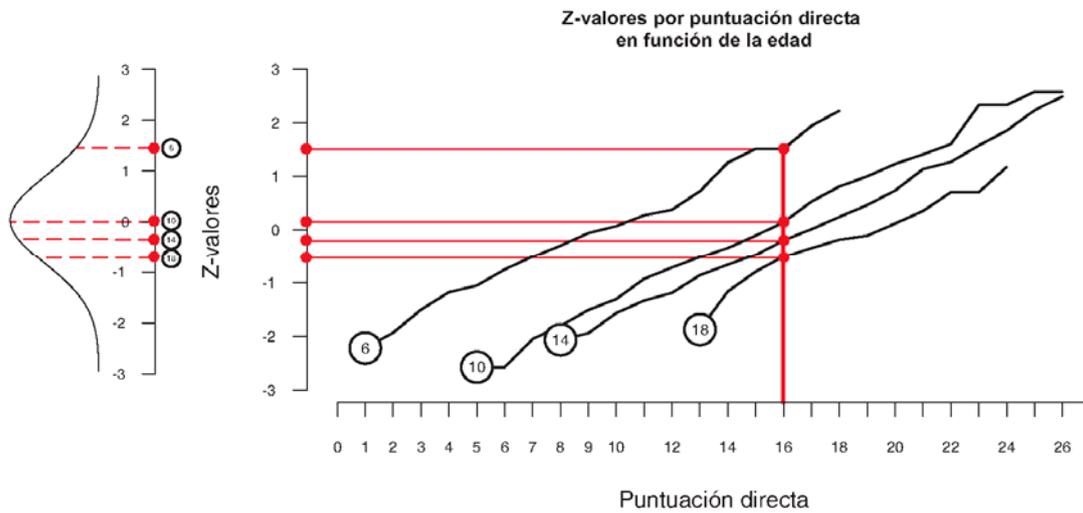
Para cada grupo de edad específico, los valores Z son monótonamente ascendentes con la puntuación directa (Figuras 13, 14 y 15).

Para cada puntuación directa obtenida, los valores Z son monótonamente descendentes con la edad (Figuras 16, 17 y 18).

Estas dos características garantizan que los baremos obtenidos por grupo de edad son consistentes ya que nos sitúan coherentemente a un sujeto con una puntuación directa y una edad frente a su misma población. En la Figura 13 podemos comprobar como un sujeto con 18 años de edad y una puntuación directa de 16 tiene un valor Z de -1, frente al valor Z de 1 que obtendría un niño de 6 años de edad y la misma puntuación. En todos los casos (Figuras 13 a 21) se observa que las líneas correspondientes a las Puntuaciones Directas de cada edad no se cruzan en ningún momento. De la misma manera, las líneas que corresponden a las puntuaciones directas según la edad tampoco se cruzan.

En las siguientes figuras se muestra los valores Z para cada prueba.

Figura 13 – Valores Z por grupo de edad como función de la Puntuación Directa en la Prueba Semejanzas



A una misma puntuación directa (en la figura éste valor es 16), diferentes sujetos con diferente edad, muestran valores Z tipificados diferentes. Así, un sujeto de 6 años, obtiene un valor aproximado Z de 1,5 (centil 93 respecto de la población de 6 años). Un sujeto de 10 años, obtiene un valor aproximado Z de 0 (centil 50 respecto de la población de 10 años). Un sujeto de 14 años, obtiene un valor aproximado Z de -0,3 (centil 38 respecto de la población de 14 años). Un sujeto de 18 años, obtiene un valor aproximado Z de -0,7 (centil 24 respecto de la población de 18 años).

Figura 14 – Valores Z por grupo de edad como función de la Puntuación Directa en la Prueba Analogías

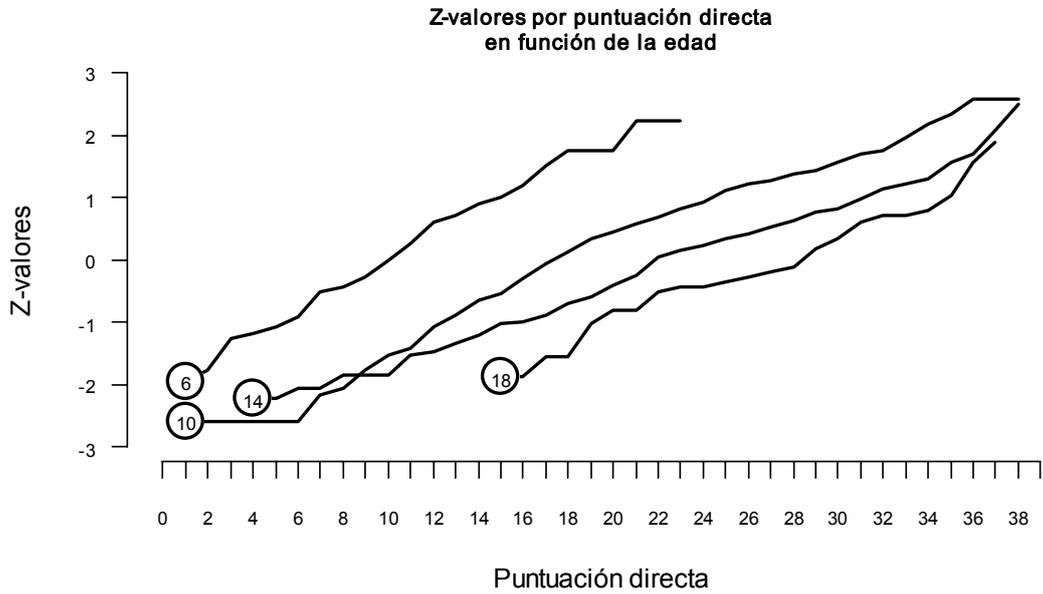


Figura 15 – Valores Z por grupo de edad como función de la Puntuación Directa en la Prueba Contextos

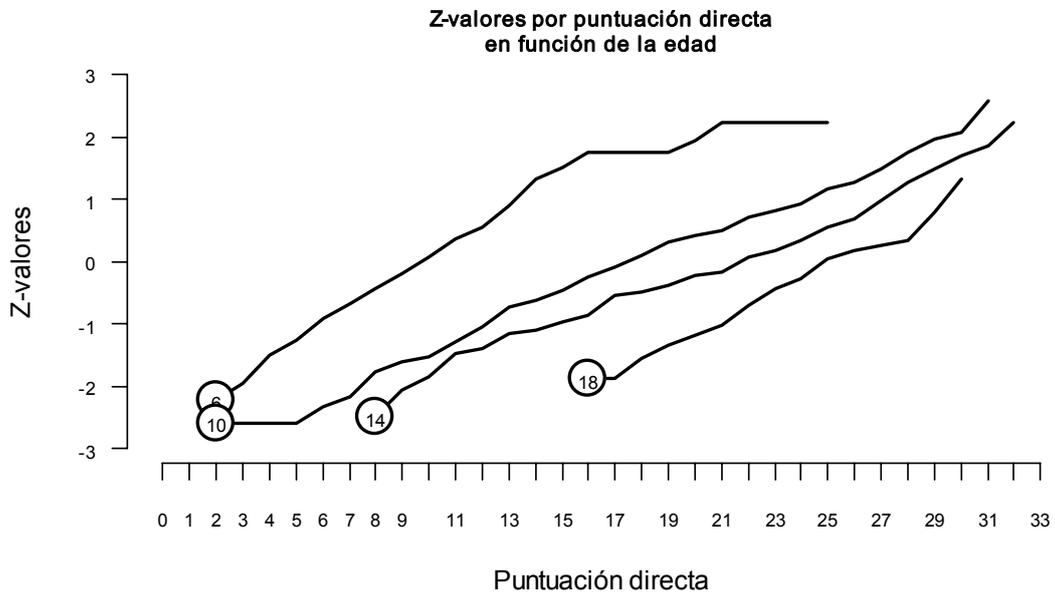


Figura 16 – Valores Z por Puntuación Directa como función de la edad en la Prueba Semejanzas

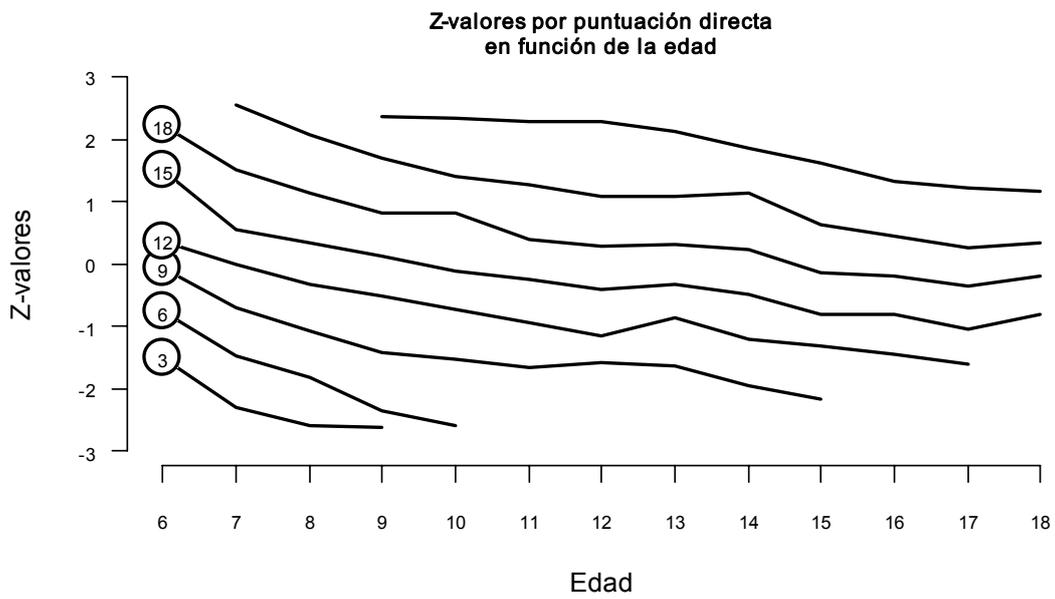


Figura 17 – Valores Z por Puntuación Directa como función de la edad en la Prueba Analogías

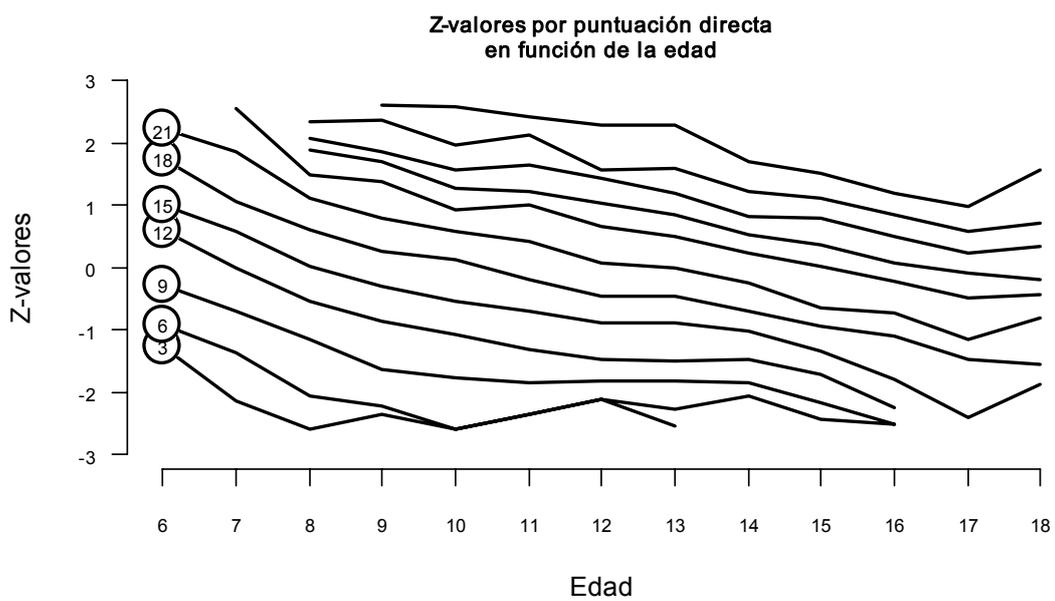
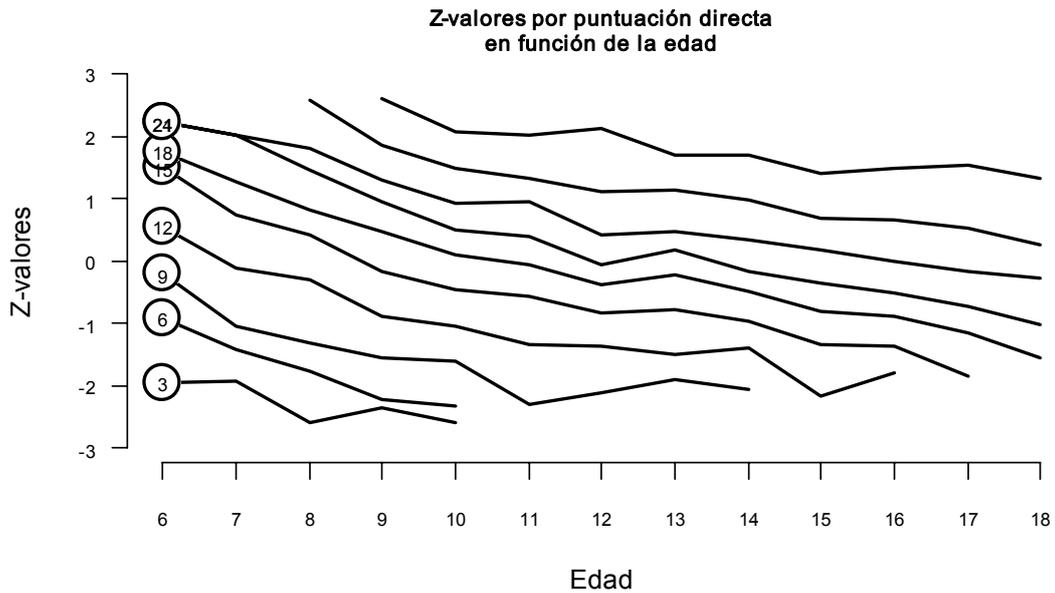


Figura 18 – Valores Z por Puntuación Directa como función de la edad en la Prueba Contextos



A partir de estos valores tipificados se obtiene la puntuación CI esperada para cada sujeto, expresada en forma normalizada. Así, un sujeto que haya obtenido un valor Z igual a 0, se le asociará un CI 100.

En las siguientes figuras se muestra los valores CI para cada prueba, según edad y puntuación directa obtenida en cada prueba.

Figura 19 - CI grupo de edad como función de la Puntuación Directa en la Prueba Semejanzas

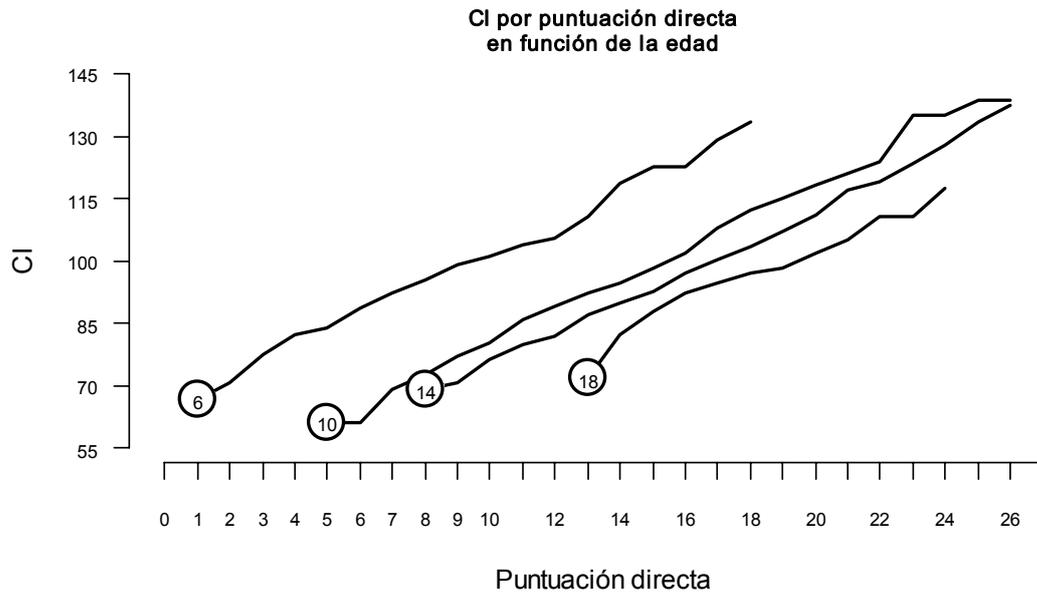


Figura 20 - CI grupo de edad como función de la Puntuación Directa en la Prueba Analogías

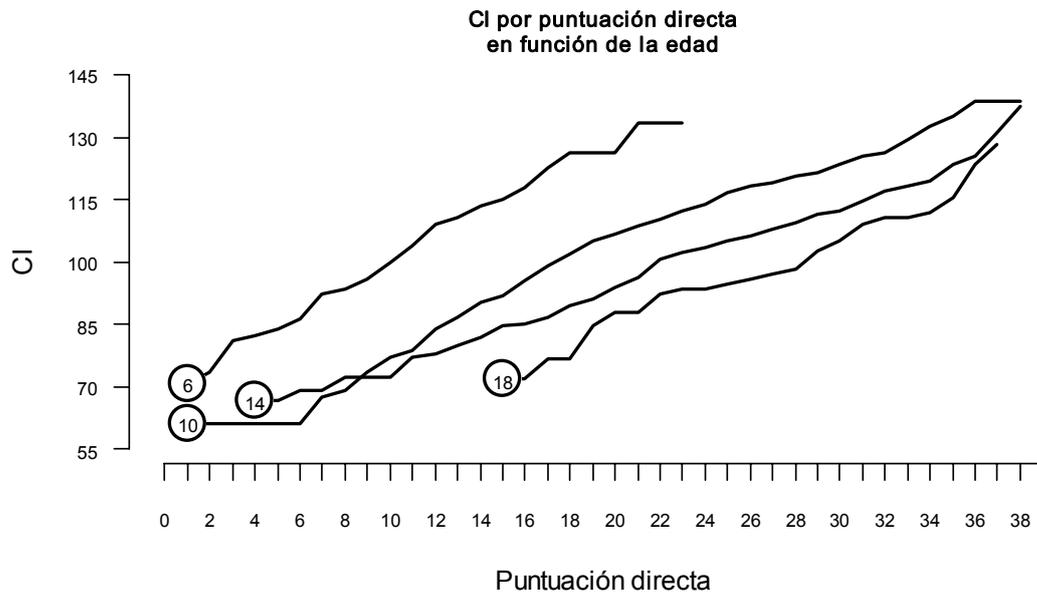
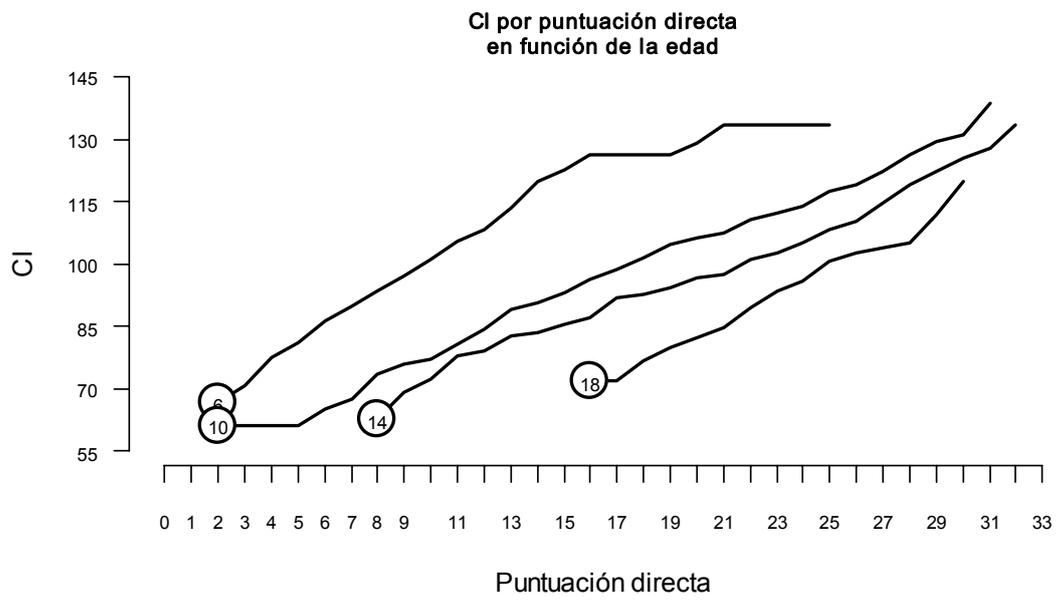


Figura 21 - CI grupo de edad como función de la Puntuación Directa en la Prueba Contextos



4.1. Homogeneidad de los perfiles

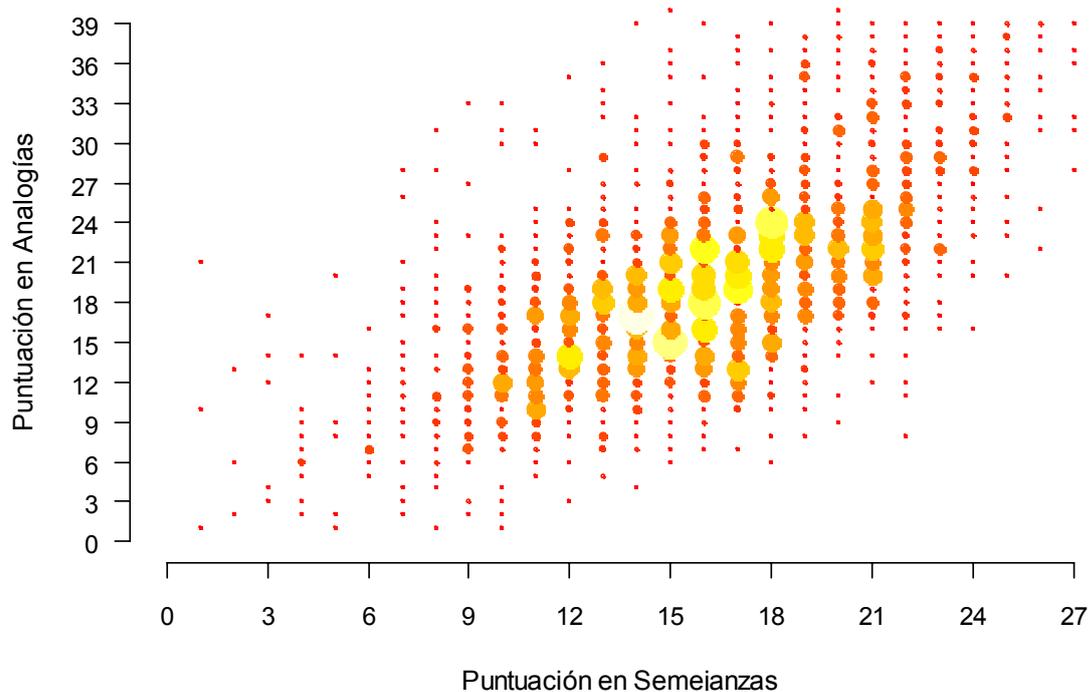
Como ya hemos definido en nuestro modelo para el cálculo de CI, dependemos de las tres puntuaciones directas obtenidas en las pruebas semejanzas, analogías y contextos.

Las puntuaciones directas obtenidas en las tres pruebas presentan las siguientes correlaciones:

Tabla 47 Correlación entre pruebas

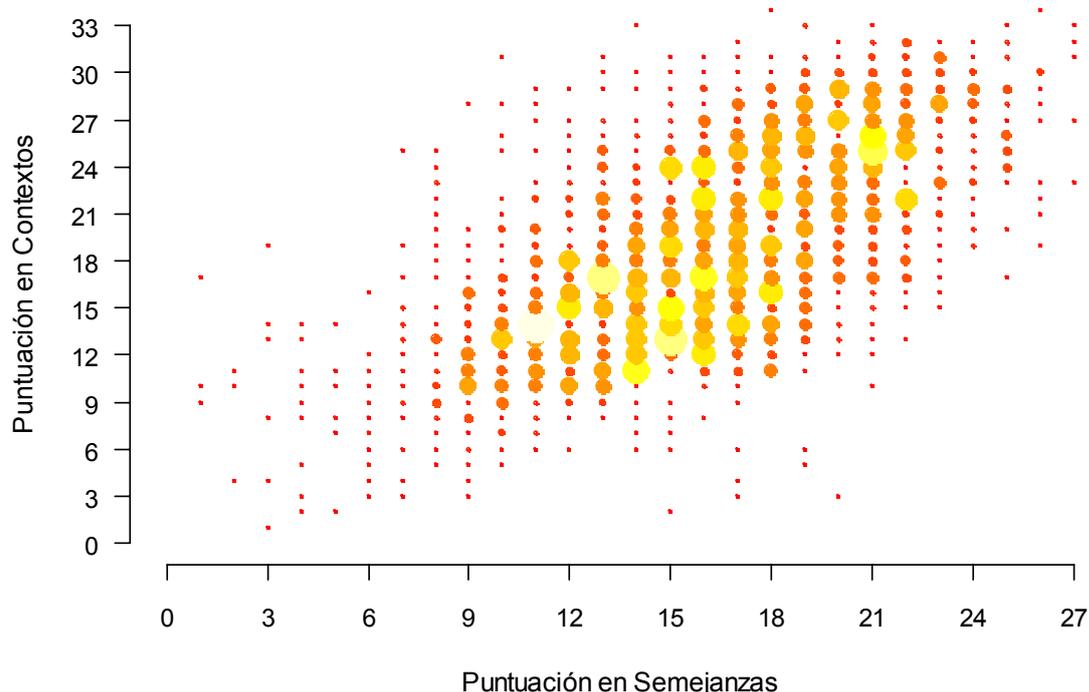
	Semejanzas	Analogías
Analogías	0,64	-
Contextos	0,63	0,72

Estas altas correlaciones en las pruebas nos han hecho plantearnos el concepto de perfil homogéneo y perfil heterogéneo a partir del comportamiento de las puntuaciones directas obtenidas en las tres pruebas. En la siguiente figura mostramos cómo se han comportando las puntuaciones directas obtenidas en la prueba Semejanzas y la prueba Analogías.

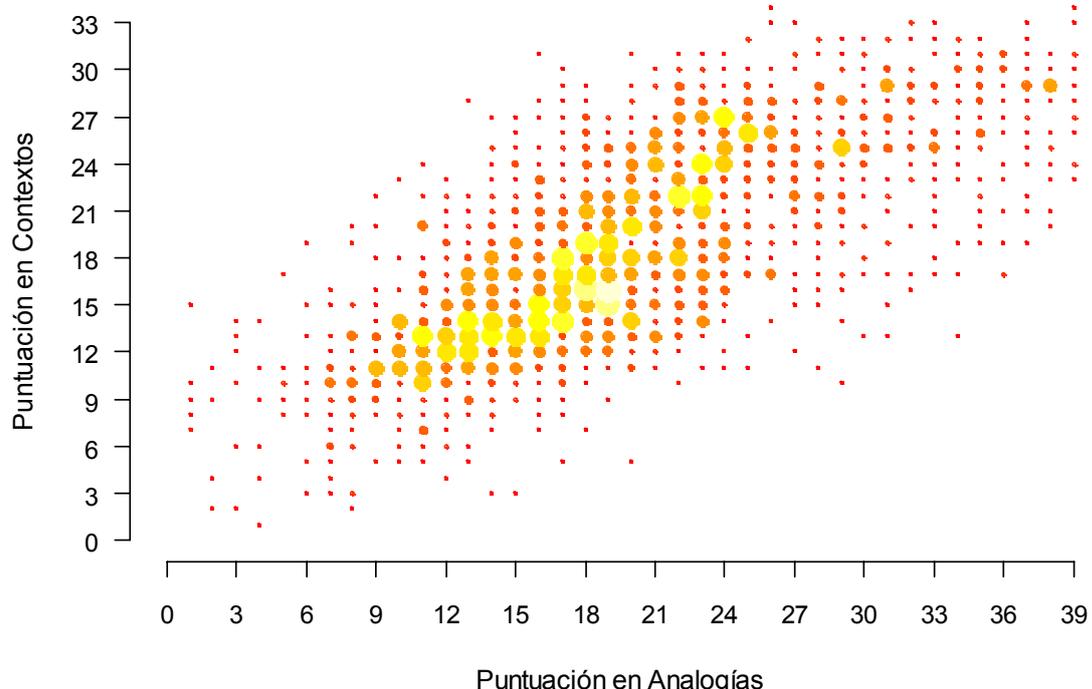


El tamaño y color de los puntos de la nube de puntos representa el número de sujetos que han obtenido la correspondiente puntuación directa en semejanzas y analogías. A mayor tamaño más sujetos en la misma situación. El color rojo implica pocos sujetos y el color amarillo-blanco implica muchos sujetos.

En la siguiente figura mostramos cómo se han comportando las puntuaciones directas obtenidas en la prueba Semejanzas y la prueba Contextos.



En la siguiente figura mostramos como se han comportando las puntuaciones directas obtenidas en la prueba Analogías y la prueba Contextos.



A partir de estos datos, hemos definido las siguientes probabilidades condicionadas:

$$E[\text{Analogías} \in (\text{PD_A}_1, \text{PD_A}_2) \mid \text{Semejanzas} = \text{PD_S}] \quad [1.1]$$

Que debe entenderse como la esperanza de obtener una puntuación directa en la prueba Analogías entre PD_A_1 e PD_A_2 , habiendo obtenido una puntuación directa PD_S en la prueba semejanzas. Equivalentemente, definimos las siguientes probabilidades condicionadas

$$E[\text{Contextos} \in (\text{PD_C}_1, \text{PD_C}_2) \mid \text{Semejanzas} = \text{PD_S}] \quad [1.2]$$

$$E[\text{Contextos} \in (\text{PD_C}_1, \text{PD_C}_2) \mid \text{Analogías} = \text{PD_A}] \quad [1.3]$$

Los valores PD_C_1 y PD_C_2 de [1.2] y [1.3] no tienen porque coincidir.

En esta situación, podemos precisar los valores PD_{A_1} e PD_{A_2} (y de igual modo PD_{C_1} , PD_{C_2}) para definir un intervalo que cumpla:

$$E[\text{Analogías} \in (PD_{A_1}, PD_{A_2}) \mid \text{Semejanzas} = PD_S] = 0.5$$

$$E[\text{Analogías} < PD_{A_1} \mid \text{Semejanzas} = PD_S] = 0.25$$

$$E[\text{Analogías} < PD_{A_2} \mid \text{Semejanzas} = PD_S] = 0.75$$

Estas esperanzas las obtenemos muestralmente, y también los valores PD_{A_1} , PD_{A_2} , PD_{C_1} , PD_{C_2} . Con ellas definimos los conceptos de perfil homogéneo y perfil heterogéneo.

Dado un sujeto con puntuaciones directas PS , PA y PC en semejanzas, analogías y contextos respectivamente, se dice que:

Tiene un perfil homogéneo cuando se cumplen las siguientes condiciones:

$$PA \in (PD_{A_1}, PD_{A_2})$$

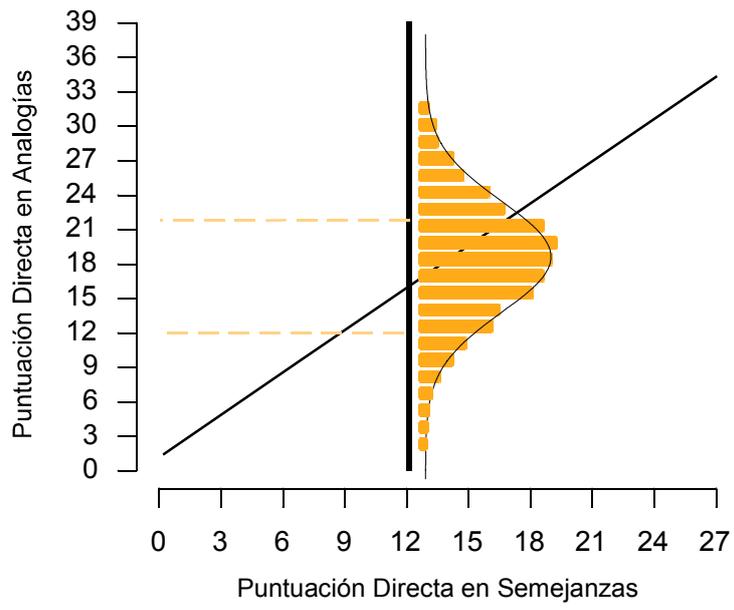
$$PC \in (PD_{C_1}, PD_{C_2})$$

Tiene un perfil heterogéneo cuando no se cumple alguna de las siguientes condiciones:

$$PA \in (Y_1, Y_2)$$

$$PC \in (X_1, X_2)$$

$$PC \in (W_1, W_2)$$



4.2. Intervalos para la puntuación verdadera

A la puntuación obtenida en cada prueba hemos asociado un intervalo de confianza, variable en cada edad y prueba, que depende del Error Típico de Medida. De esta manera, hemos determinado unos límites dentro de los cuales se encuentra la puntuación verdadera de un sujeto, asumiendo igualdad de las varianzas para todas las puntuaciones directas obtenidas para cada grupo.

Hemos estimado el Error Máximo Admisible de la siguiente manera:

$$E_{m\acute{a}x} = |z_c| \cdot (\hat{\sigma}_x \cdot \sqrt{1 - \hat{\rho}_{xx'}}) \quad [1.4]$$

Donde hemos considerado:

- Z_c : El valor crítico de la distribución normal estandarizada.
- $\hat{\sigma}_x$: El estimador muestral de la desviación típica de las puntuaciones observadas
- $\hat{\rho}_{xx'}$: El estimador del coeficiente de fiabilidad

Aunque este intervalo se construya sobre la puntuación directa obtenida, las propiedades quedan invariantes si realizamos cambios de origen y de unidades de la escala, por lo que extrapolamos estos intervalos a intervalos de confianza para el CI estimado.

Por todo esto, los intervalos para el CI estimado, no serán de la misma longitud si consideramos sujetos de distinta edad, ya que hemos observado distintas varianzas muestrales y diferentes coeficientes de fiabilidad en cada grupo.

5. Propiedades Psicométricas

La fiabilidad de la **Batería CERVANTES** se ha evaluado en términos de consistencia interna y fiabilidad test-retest, que tiene en cuenta la varianza de las observaciones con el tiempo.

Se han estudiado las siguientes características psicométricas de la **Batería CERVANTES**:

Validez

Fiabilidad de las series

5.1. Validez del constructo, concurrente y discriminante

5.1.1. Test-retest

La estimación del coeficiente de fiabilidad por el procedimiento test-retest requiere aplicar la **Batería CERVANTES** dos veces. La correlación entre las puntuaciones de las dos ocasiones, será un estimador de la fiabilidad. La aplicación de esta prueba se ha realizado a 35 sujetos de entre 7 y 12 años de edad, con un intervalo de tiempo entre aplicaciones de 8 semanas.

Para no sobreestimar ni infravalorar este estimador, se ha tenido en cuenta las siguientes recomendaciones a la hora de realizar la segunda prueba:

1. A los sujetos que han participado en el test-retest no se les ha informado sobre su aciertos y fallos en las pruebas en la primera aplicación, para no influir en las puntuaciones de la segunda aplicación.
2. El intervalo de tiempo entre las dos administraciones de la **Batería CERVANTES** puede afectar también al cálculo del estimador, pero dado que el factor memoria no afecta a las pruebas de la **Batería CERVANTES**, los investigadores consideraron oportuno un intervalo de 8 semanas.

La correlación entre las puntuaciones obtenidas en las dos aplicaciones resultó ser 0,87.

5.1.2. Validez concurrente y validez discriminante

Nos referiremos a validez concurrente al grado en el que las puntuaciones de la **Batería CERVANTES** correlacionan con otras pruebas que evalúan el criterio de inteligencia definido por los investigadores, considerando la validez concurrente como un indicador de que el uso de la **Batería CERVANTES** es adecuado.

Para el cálculo de la validez concurrente, se han aplicado a 50 sujetos la **Batería CERVANTES**, el Test de Raven, las pruebas de razonamiento de BADYG-R y el SON-R, obteniéndose los siguientes resultados:

Tabla 48 Correlación entre la Batería CERVANTES y otras pruebas de evaluación de la Inteligencia

	Test de Raven	Pruebas de Razonamiento de BADYG-R*	SON-R**
Batería CERVANTES	0,70	0,75	0,80

* En las Escalas E1, E2, E3, M y S y en las pruebas Analogías Verbales, Series numéricas y Matrices de Figuras

** En las pruebas

Observamos que hay altas correlaciones entre las pruebas, siendo la más alta 0,80 entre la **Batería CERVANTES** y el SON-R.

La validez discriminante se definirá como el grado en el que las puntuaciones de la **Batería CERVANTES** correlacionan con otros criterios. Es esperable que la correlación de las puntuaciones de la **Batería CERVANTES** con otras pruebas que no midan inteligencia, sea baja o incluso negativa. Para el cálculo de la validez discriminante se han aplicado las pruebas de Comprensión Lectora del TALE-2000, la Escala de Atención Sostenida de la EMAY-1 y la Escala de Calidad Atencional de la EMAY-1, obteniéndose los siguientes resultados:

Tabla 49 Estudio de la validez discriminante de la Batería CERVANTES

	Comprensión Lectora TALE-2000	Escala de Atención Sostenida EMAY-1	Escala de Calidad Atencional EMAY-1
Batería CERVANTES	0,57	0,20	0,65

5.2. Fiabilidad

La fiabilidad de cada serie se basa en la consistencia interna de cada prueba. Como las pruebas están formadas por un grupo de preguntas homogéneas garantizado en gran medida por la aplicación del modelo de Mokken, la consistencia interna da una estimación precisa de la fiabilidad. Para calcular la fiabilidad hemos adoptado el Coeficiente Alpha de Cronbach, que viene dado por la ecuación:

$$\alpha = \frac{n}{n-1} \left(1 - \frac{\sum_{j=1}^n \sigma_j^2}{\sigma_x^2} \right) \quad [1.5]$$

Donde:

- n : Es el número de elementos de la Prueba
- σ_j^2 : Es la varianza de un ítem ($j = 1, \dots, n$)
- σ_x^2 : Es la varianza de la Prueba

Las varianzas se han considerado mediante estimadores muestrales. Este procedimiento del cálculo de fiabilidad está basado en el supuesto de que las puntuaciones de las partes son esencialmente tau-equivalentes, tal como se demuestra en los siguientes apartados, de igualdad de medias y de varianzas.

En los siguientes apartados, consideramos los siguientes ítems:

Parte A: {Ítem 1-Serie A, Ítem 1-Serie C, Ítem 2-Serie B, Ítem 3-Serie A, ...}

Parte B: {Ítem 1-Serie B, Ítem 2-Serie A, Ítem 2-Serie C, Ítem 3-Serie B, ...}

5.2.1. Igualdad de medias

El primer supuesto de paralelismo implica la igualdad de las puntuaciones (Media (Parte A) = Media (Parte B)), lo cual puede observarse en el siguiente gráfico en el que se muestran la media y el intervalo de confianza al 95% (obtenido mediante una prueba no paramétrica), para cada prueba de la Batería CERVANTES por grupo de edad.

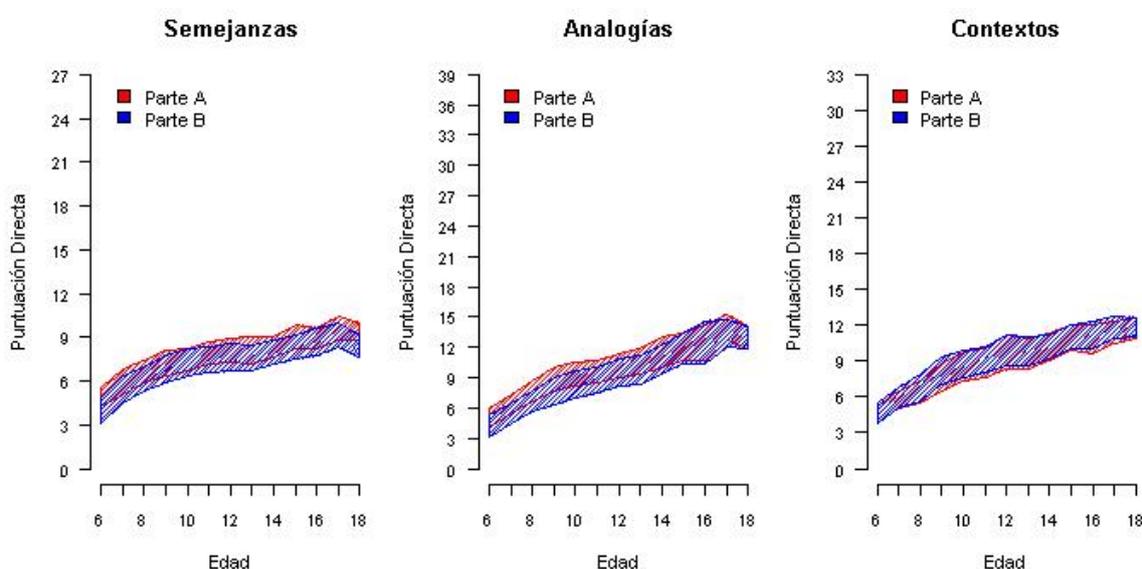


Tabla 50 Distribución de la puntuación media de cada parte de la prueba y significación

	Semejanzas			Analogías			Contextos		
	Parte A	Parte B	Significación	Parte A	Parte B	Significación	Parte A	Parte B	Significación
6 Años	5,01	4,11	0,31	5,14	4,33	0,41	4,73	4,76	0,73
7 Años	6,02	5,47	0,49	6,33	5,35	0,31	5,81	5,83	0,69
8 Años	6,61	6,15	0,55	7,67	6,81	0,44	6,46	6,71	0,69
9 Años	7,06	6,65	0,60	8,79	7,73	0,35	7,74	8,11	0,66
10 Años	7,46	7,18	0,66	9,42	8,45	0,43	8,56	8,80	0,73
11 Años	7,85	7,57	0,65	9,59	8,70	0,42	8,93	9,18	0,73
12 Años	8,12	7,82	0,62	10,28	9,63	0,59	9,85	10,01	0,77
13 Años	8,09	7,64	0,58	10,68	9,87	0,53	9,65	9,88	0,76
14 Años	8,28	7,90	0,59	11,46	10,82	0,62	10,16	10,14	0,79
15 Años	9,04	8,46	0,46	11,93	11,67	0,79	10,90	11,09	0,72
16 Años	8,97	8,83	0,72	12,70	12,35	0,72	11,08	11,34	0,69
17 Años	9,51	9,10	0,57	13,80	13,33	0,66	11,55	11,93	0,58
18 Años	9,49	8,57	0,25	13,13	13,01	0,82	11,69	11,87	0,74

5.2.2. Igualdad de varianzas

El segundo supuesto de paralelismo implica la igualdad de las varianzas, lo cual puede observarse en el siguiente gráfico en el que se muestra el intervalo de confianza al 95% para la varianza por grupo de edad (obtenido mediante una prueba no paramétrica), para cada prueba de la **Batería CERVANTES** por grupo de edad.

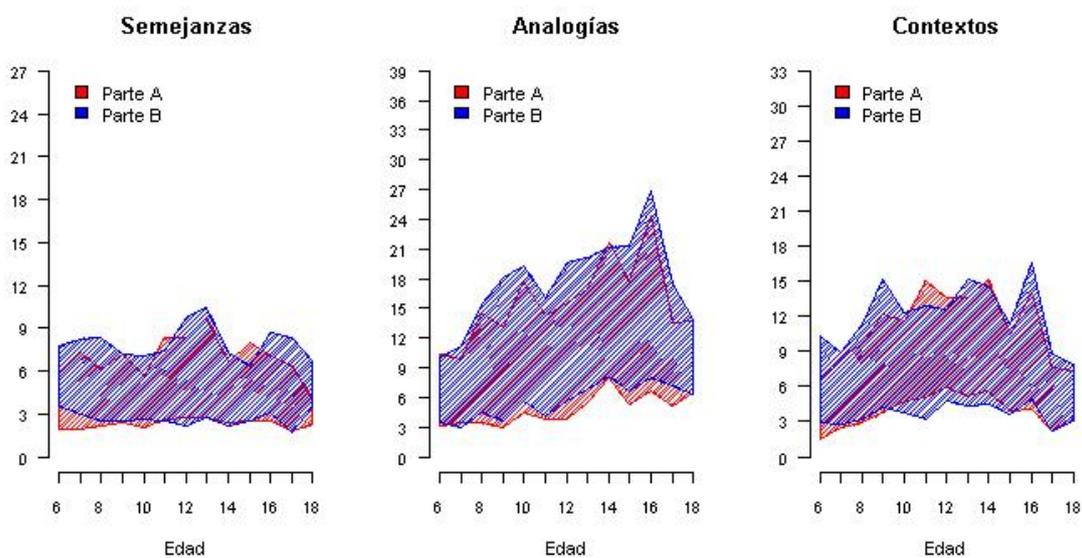


Tabla 51 Distribución de la varianza de cada parte de la prueba y significación

	Semejanzas			Analogías			Contextos		
	Parte A	Parte B	Significación	Parte A	Parte B	Significación	Parte A	Parte B	Significación
Edad 6	4,16	6,02	0,49	5,65	6,26	0,63	3,48	5,23	0,45
Edad 7	4,53	5,50	0,62	5,96	6,73	0,70	5,16	5,65	0,66
Edad 8	3,88	5,58	0,50	7,67	9,26	0,63	5,29	6,59	0,63
Edad 9	4,20	4,67	0,66	7,44	10,01	0,57	7,68	7,75	0,67
Edad 10	3,77	4,86	0,57	8,28	10,24	0,61	8,36	8,38	0,69
Edad 11	5,25	4,70	0,70	8,50	9,70	0,69	9,46	8,51	0,67
Edad 12	4,92	5,03	0,67	10,28	13,64	0,57	9,85	8,68	0,73
Edad 13	5,55	5,27	0,71	9,55	12,70	0,58	8,82	8,57	0,69
Edad 14	3,86	4,30	0,67	12,70	13,85	0,74	10,09	8,37	0,66
Edad 15	4,71	4,39	0,66	10,31	13,36	0,62	6,85	7,00	0,70
Edad 16	4,11	5,23	0,58	11,51	15,33	0,62	6,42	7,72	0,63
Edad 17	4,19	4,88	0,66	9,63	11,34	0,70	4,61	5,09	0,58
Edad 18	3,25	4,92	0,47	10,04	10,05	0,76	5,82	5,52	0,66

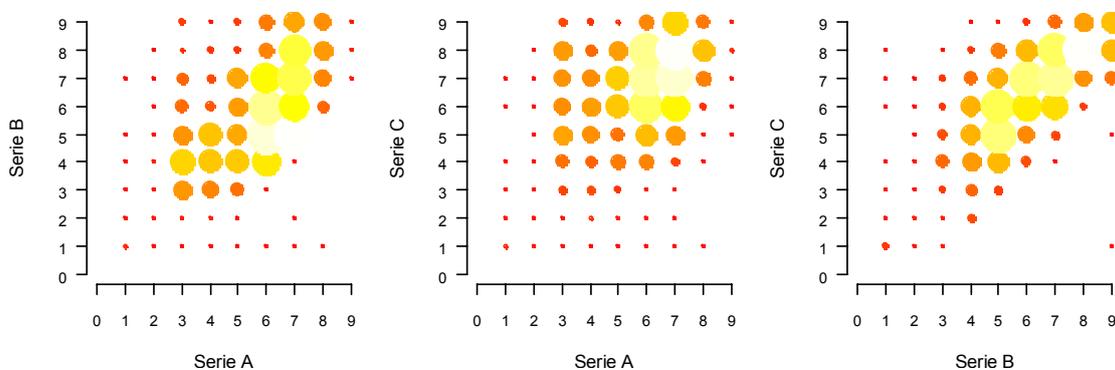
Aceptamos que ambas partes de cada Prueba de la **Batería CERVANTES** son tau-equivalentes.

Tabla 52. Correlación entre series según prueba

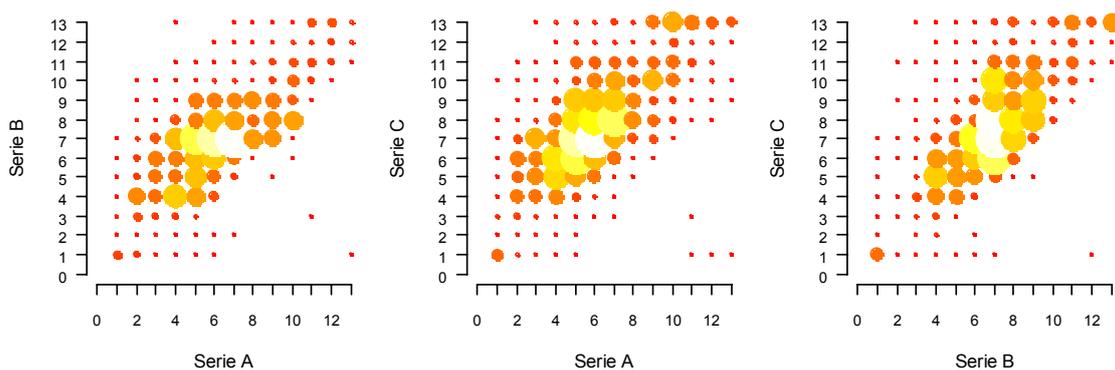
	Semejanzas		Analogías		Contextos	
	Serie A	Serie B	Serie A	Serie B	Serie A	Serie B
Serie B	0,52	-	0,77	-	0,68	-
Serie C	0,39	0,71	0,71	0,77	0,64	0,83

Todas las correlaciones entre series son significativamente diferentes de 0.

Distribución de puntuaciones entre las series de la Prueba Semejanzas



Distribución de puntuaciones entre las series de la Prueba Analogías



Distribución de puntuaciones entre las series de la Prueba Contextos

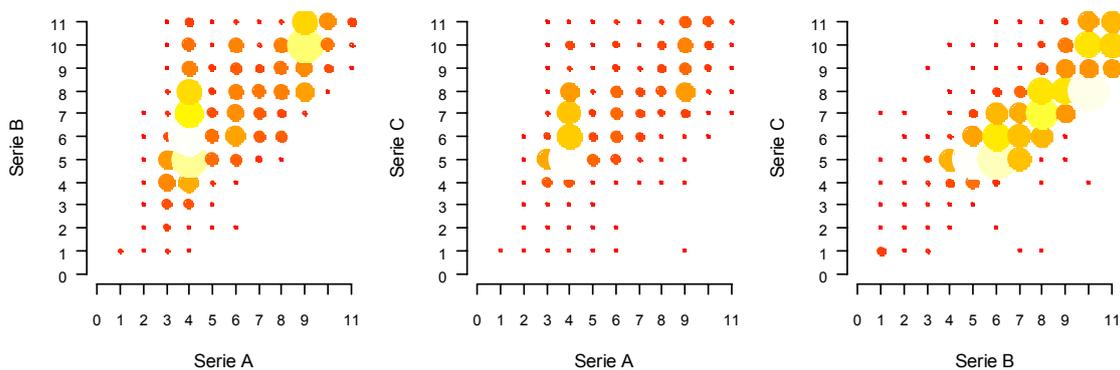


Tabla 53 Correlación entre pruebas según edad

Edad	Semejanzas-Analogías	Semejanzas-Contextos	Analogías-Contextos		Media
6	0,47	0,51	0,60		0,53
7	0,52	0,50	0,57		0,53
8	0,54	0,50	0,58		0,54
9	0,46	0,46	0,59		0,50
10	0,57	0,49	0,59		0,55
11	0,44	0,44	0,50		0,46
12	0,59	0,52	0,63		0,58
13	0,49	0,52	0,60		0,54
14	0,51	0,52	0,60		0,54
15	0,48	0,38	0,56		0,47
16	0,45	0,47	0,57		0,50
17	0,40	0,44	0,40		0,41
18	0,59	0,47	0,69		0,58
Media	0,50	0,48	0,58		0,52

En las siguientes tablas se describe para cada ítem: el índice de dificultad, definido como la proporción de sujetos que ha acertado el ítem, la desviación típica y la correlación del ítem con el total de la serie.

Tabla 54 Índice de Dificultad, Desviación Típica y Correlación con la serie, de los ítems de la Serie A de la prueba Semejanzas.

	ID	Desviación Típica	Correlación con el Total
Ítem 1	0,98	0,16	0,21
Ítem 2	0,96	0,20	0,32
Ítem 3	0,65	0,48	0,27
Ítem 4	0,62	0,49	0,35
Ítem 5	0,73	0,45	0,47
Ítem 6	0,67	0,47	0,39
Ítem 7	0,40	0,49	0,43
Ítem 8	0,12	0,33	0,38
Ítem 9	0,10	0,29	0,28

Tabla 55 Índice de Dificultad, Desviación Típica y Correlación con la serie, de los ítems de la Serie B de la prueba Semejanzas.

	ID	Desviación Típica	Correlación con el Total
Ítem 1	0,98	0,13	0,25
Ítem 2	0,98	0,15	0,28
Ítem 3	0,84	0,36	0,54
Ítem 4	0,80	0,40	0,52
Ítem 5	0,54	0,50	0,60
Ítem 6	0,58	0,49	0,53
Ítem 7	0,53	0,50	0,46
Ítem 8	0,39	0,49	0,47
Ítem 9	0,33	0,47	0,32

Tabla 56 Índice de Dificultad, Desviación Típica y Correlación con la serie, de los ítems de la Serie C de la prueba Semejanzas.

	ID	Desviación Típica	Correlación con el Total
Ítem 1	0,98	0,13	0,28
Ítem 2	0,94	0,24	0,44
Ítem 3	0,92	0,27	0,40
Ítem 4	0,84	0,37	0,47
Ítem 5	0,70	0,46	0,54
Ítem 6	0,69	0,46	0,43
Ítem 7	0,50	0,50	0,46
Ítem 8	0,36	0,48	0,34
Ítem 9	0,20	0,40	0,44

Tabla 57 Índice de Dificultad, Desviación Típica y Correlación con la serie, de los ítems de la Serie A de la prueba Analogías.

	ID	Desviación Típica	Correlación con el Total
Ítem 1	0,98	0,14	0,21
Ítem 2	0,60	0,49	0,31
Ítem 3	0,90	0,31	0,28
Ítem 4	0,84	0,36	0,25
Ítem 5	0,77	0,42	0,28
Ítem 6	0,61	0,49	0,22
Ítem 7	0,59	0,49	0,29
Ítem 8	0,59	0,49	0,22
Ítem 9	0,70	0,46	0,16
Ítem 10	0,48	0,50	0,17
Ítem 11	0,44	0,50	0,18
Ítem 12	0,49	0,50	0,12
Ítem 13	0,33	0,47	0,09

Tabla 58 Índice de Dificultad, Desviación Típica y Correlación con la serie, de los ítems de la Serie B de la prueba Analogías.

	ID	Desviación Típica	Correlación con el Total
Ítem 1	0,98	0,13	0,19
Ítem 2	0,98	0,13	0,16
Ítem 3	0,95	0,23	0,26
Ítem 4	0,83	0,37	0,33
Ítem 5	0,72	0,45	0,38
Ítem 6	0,78	0,41	0,25
Ítem 7	0,66	0,47	0,28
Ítem 8	0,35	0,48	0,25
Ítem 9	0,57	0,49	0,26
Ítem 10	0,51	0,50	0,27
Ítem 11	0,55	0,50	0,12
Ítem 12	0,44	0,50	0,18
Ítem 13	0,40	0,49	0,11

Tabla 59 Índice de Dificultad, Desviación Típica y Correlación con la serie, de los ítems de la Serie C de la prueba Analogías.

	ID	Desviación Típica	Correlación con el Total
Ítem 1	0,99	0,10	0,18
Ítem 2	0,98	0,15	0,24
Ítem 3	0,95	0,21	0,26
Ítem 4	0,92	0,28	0,29
Ítem 5	0,85	0,36	0,37
Ítem 6	0,73	0,50	0,28
Ítem 7	0,70	0,46	0,33
Ítem 8	0,62	0,49	0,32
Ítem 9	0,54	0,50	0,23
Ítem 10	0,67	0,47	0,23
Ítem 11	0,55	0,50	0,17
Ítem 12	0,54	0,50	0,17
Ítem 13	0,64	0,48	0,02

Tabla 60 Índice de Dificultad, Desviación Típica y Correlación con la serie, de los ítems de la Serie A de la prueba Contextos.

	ID	Desviación Típica	Correlación con el Total
Ítem 1	1,00	0,07	0,08
Ítem 2	0,98	0,13	0,14
Ítem 3	0,64	0,48	0,30
Ítem 4	0,79	0,41	0,31
Ítem 5	0,31	0,46	0,30
Ítem 6	0,82	0,38	0,27
Ítem 7	0,66	0,47	0,09
Ítem 8	0,74	0,44	0,13
Ítem 9	0,62	0,48	0,11
Ítem 10	0,24	0,43	0,14
Ítem 11	0,19	0,39	0,11

Tabla 61 Índice de Dificultad, Desviación Típica y Correlación con la serie, de los ítems de la Serie B de la prueba Contextos.

	ID	Desviación Típica	Correlación con el Total
Ítem 1	0,96	0,20	0,23
Ítem 2	0,97	0,17	0,21
Ítem 3	0,93	0,26	0,26
Ítem 4	0,91	0,29	0,29
Ítem 5	0,81	0,40	0,34
Ítem 6	0,65	0,48	0,31
Ítem 7	0,60	0,49	0,20
Ítem 8	0,69	0,46	0,24
Ítem 9	0,58	0,49	0,26
Ítem 10	0,55	0,50	0,17
Ítem 11	0,19	0,39	0,19

Tabla 62 Índice de Dificultad, Desviación Típica y Correlación con la serie, de los ítems de la Serie C de la prueba Contextos.

	ID	Desviación Típica	Correlación con el Total
Ítem 1	0,98	0,15	0,21
Ítem 2	0,97	0,18	0,25
Ítem 3	0,98	0,15	0,17
Ítem 4	0,94	0,31	0,18
Ítem 5	0,71	0,45	0,38
Ítem 6	0,66	0,47	0,41
Ítem 7	0,68	0,47	0,30
Ítem 8	0,54	0,50	0,32
Ítem 9	0,34	0,47	0,21
Ítem 10	0,53	0,50	0,09
Ítem 11	0,34	0,47	0,17

La fiabilidad de las pruebas por grupo de edad puede verse resumida en la siguiente tabla.

Tabla 63 Índice Alpha de Cronbach para cada prueba por grupo de edad

Edad	Semejanzas	Analogías	Contextos	Media
6	0,80	0,84	0,82	0,82
7	0,78	0,78	0,78	0,78
8	0,69	0,82	0,81	0,77
9	0,69	0,82	0,78	0,76
10	0,70	0,81	0,84	0,78
11	0,69	0,77	0,77	0,74
12	0,70	0,81	0,83	0,78
13	0,74	0,83	0,85	0,81
14	0,70	0,88	0,87	0,82
15	0,69	0,84	0,82	0,78
16	0,62	0,72	0,71	0,69
17	0,79	0,81	0,72	0,78
18	0,48	0,90	0,56	0,65
Media	0,70	0,82	0,78	0,77

Donde podemos concluir:

La fiabilidad de la prueba tiene una media de 0,77. La fiabilidad varía de un mínimo de 0,48 en la prueba Semejanzas en el grupo de 18 años, a 0,90 en la prueba Analogías en el mismo grupo de edad.

La fiabilidad varía en cada prueba con la edad, aunque se mantiene en un rango de 0,65 a 0,82.

Si observamos la ecuación [1.5], para obtener valores cercanos a 1 (tal como obtenemos) es necesario que las covarianzas entre ítems sean altas, ya que de esa manera:

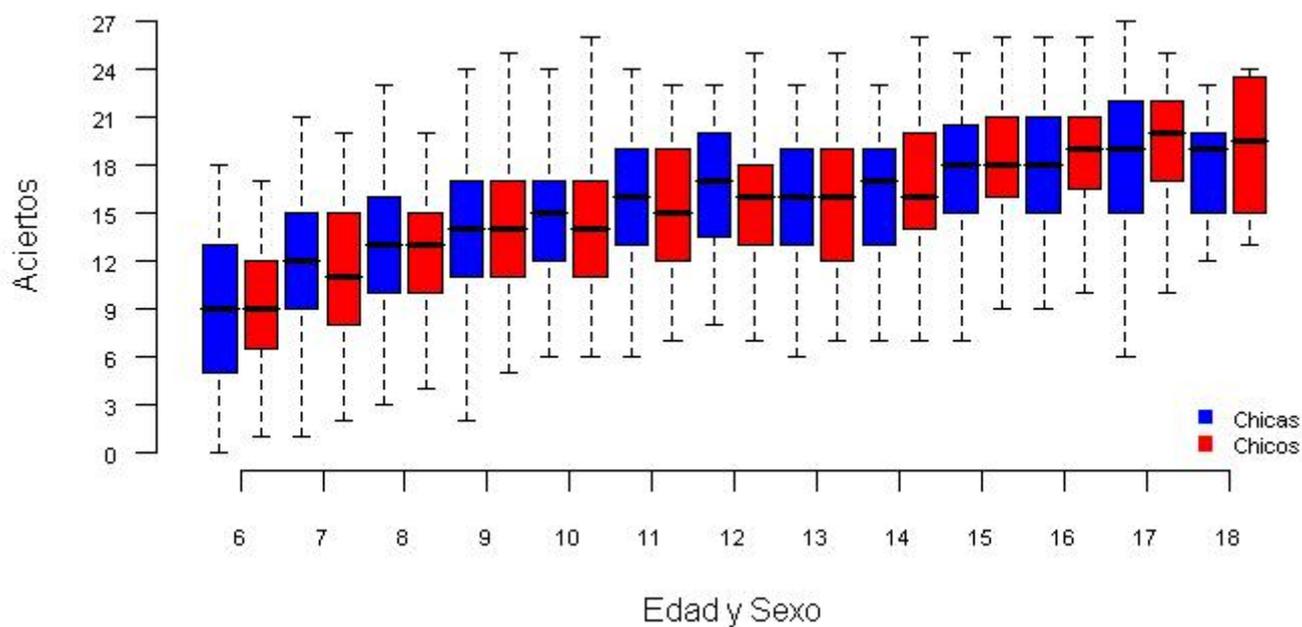
$$\frac{\sum_{j=1}^n \sigma_j^1}{\sigma_X^2} \rightarrow 0 \text{ y por tanto } \left(1 - \frac{\sum_{j=1}^n \sigma_j^2}{\sigma_X^2} \right) \rightarrow 1 \quad [1.6]$$

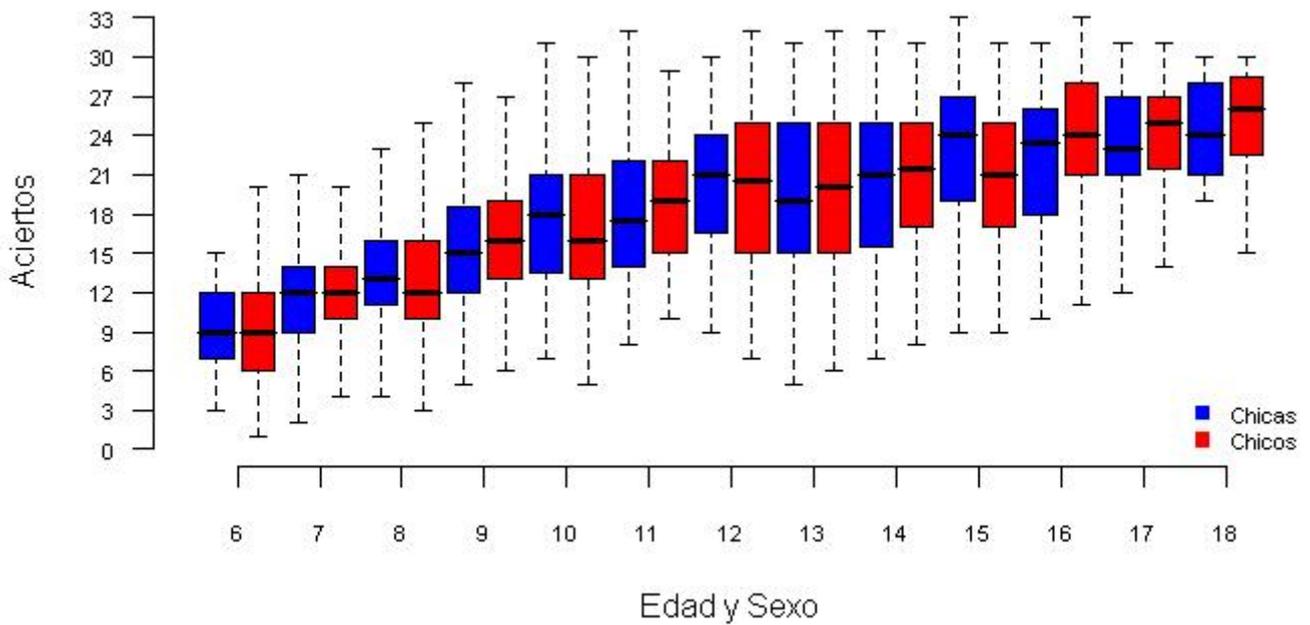
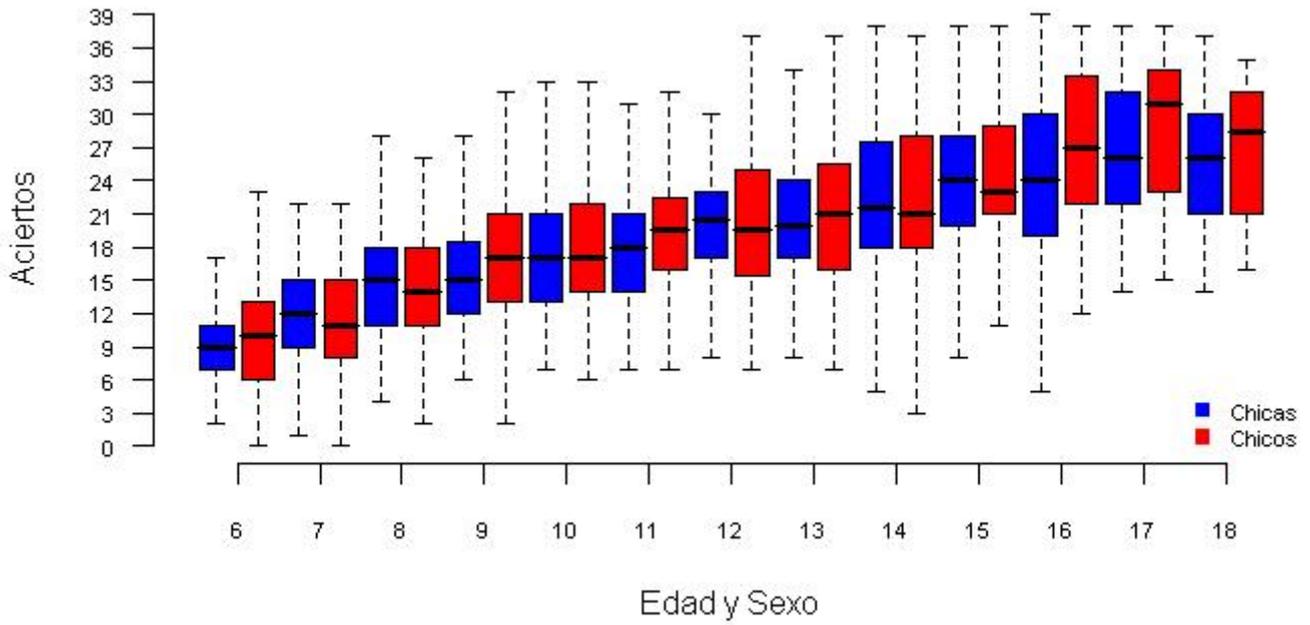
Esto es previsible, ya que la correlación entre ítems es alta y por definición, la correlación entre dos ítems, se puede calcular como la covarianza entre el producto de las varianzas.

Si la correlación entre cada par de ítems fuera 0, entonces la varianza total del test sería igual a la suma de las varianzas de los ítems y por tanto, según [1.6], el Alpha de Cronbach sería 0.

5.3. Distribución de las puntuaciones por edad y sexo

Hemos estudiado si existen diferencias significativas en las puntuaciones obtenidas por chicos y chicas en cada grupo de edad. En las siguientes figuras se muestra la distribución de las puntuaciones obtenidas por cada grupo de edad agrupando por sexo.





Para estudiar si las puntuaciones obtenidas por chicos y chicas son significativamente distintas hemos realizado la prueba de Rangos de Mann-Whitney para cada grupo de edad en cada serie. Los p-valores obtenidos en cada prueba por grupo de edad son mostrados en la siguiente tabla:

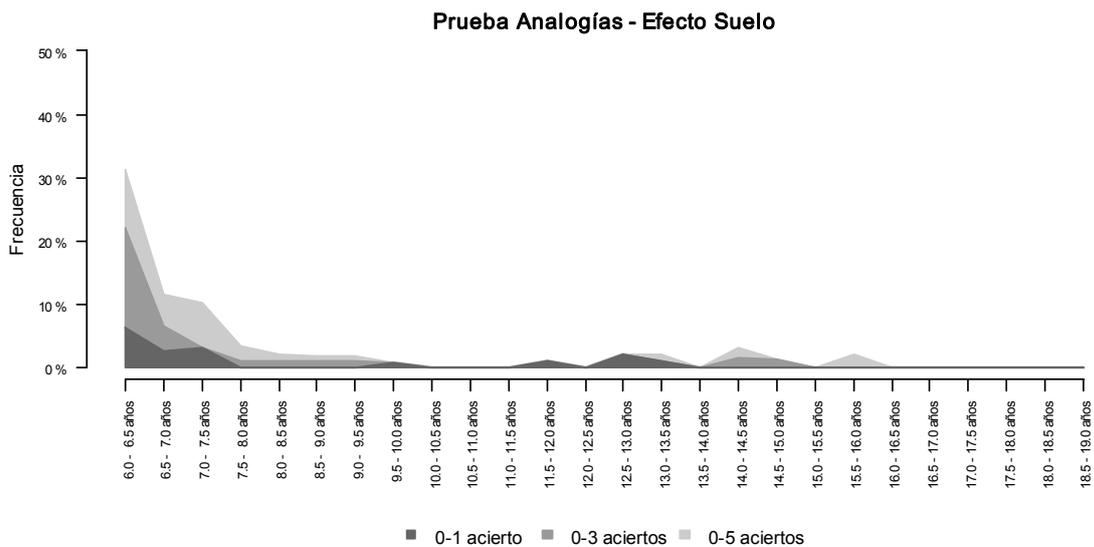
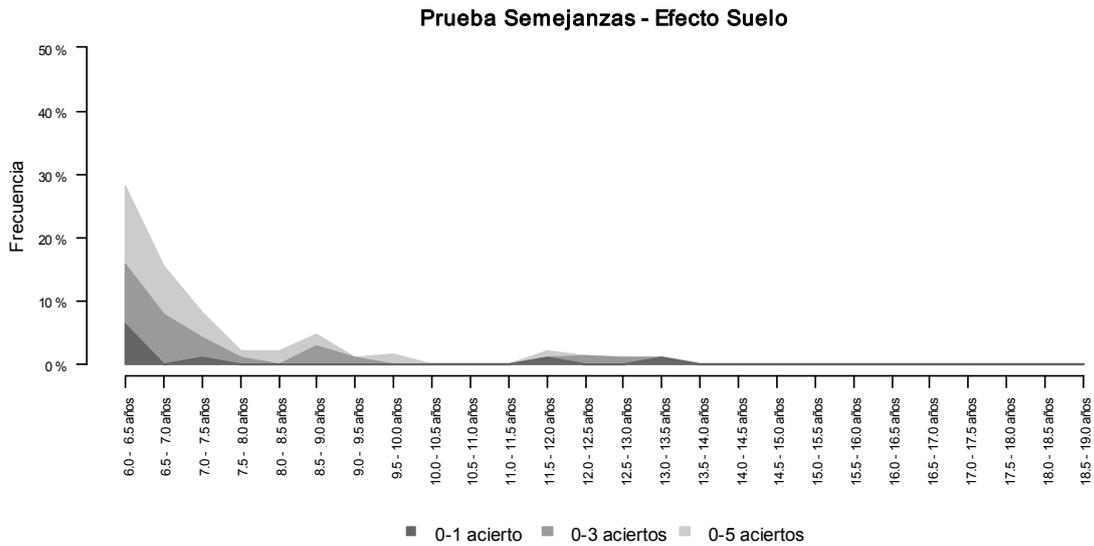
Tabla 64 Significación de la Prueba Mann-Whitney aplicada a cada Prueba por edad

Edad	Prueba		
	Semejanzas	Analogías	Contextos
6 Años	0,975	0,733	0,935
7 Años	0,428	0,431	0,807
8 Años	0,197	0,711	0,618
9 Años	0,701	0,087	0,199
10 Años	0,254	0,694	0,638
11 Años	0,533	0,046	0,274
12 Años	0,188	0,946	0,527
13 Años	0,753	0,437	0,666
14 Años	0,309	0,923	0,194
15 Años	0,482	0,675	0,046
16 Años	0,332	0,050	0,056
17 Años	0,324	0,050	0,287
18 Años	0,561	0,365	0,396

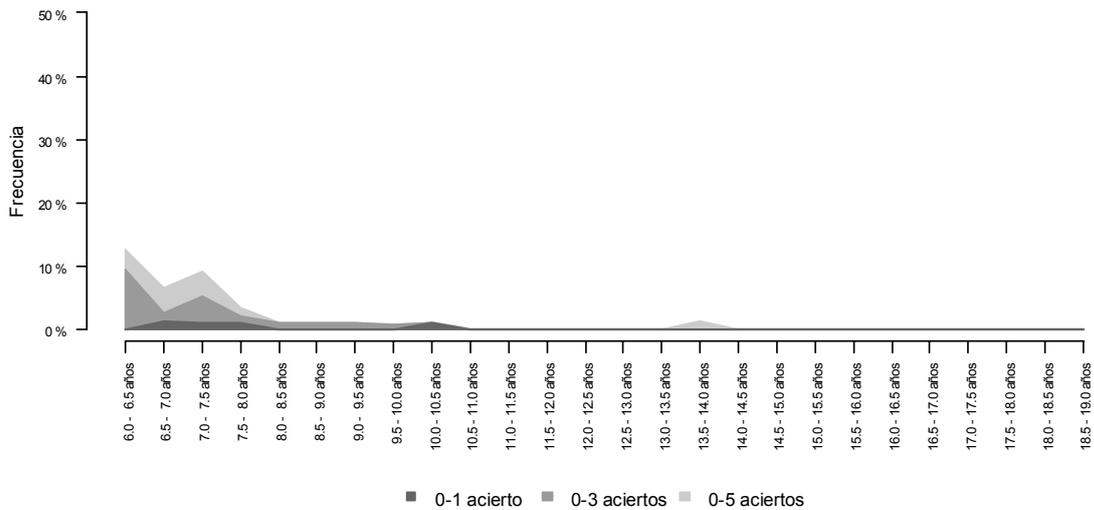
Solamente en dos grupos (Serie Analogías y 11 años y Serie Contextos y 15 años) podríamos rechazar la hipótesis nula a un nivel de significación de 0.05, es decir, que ambas distribuciones no son iguales. A partir de este punto se considerará que la distribución de las puntuaciones directas de chicos y chicas es igual en las tres pruebas.

5.4. Efectos suelo y efectos techo

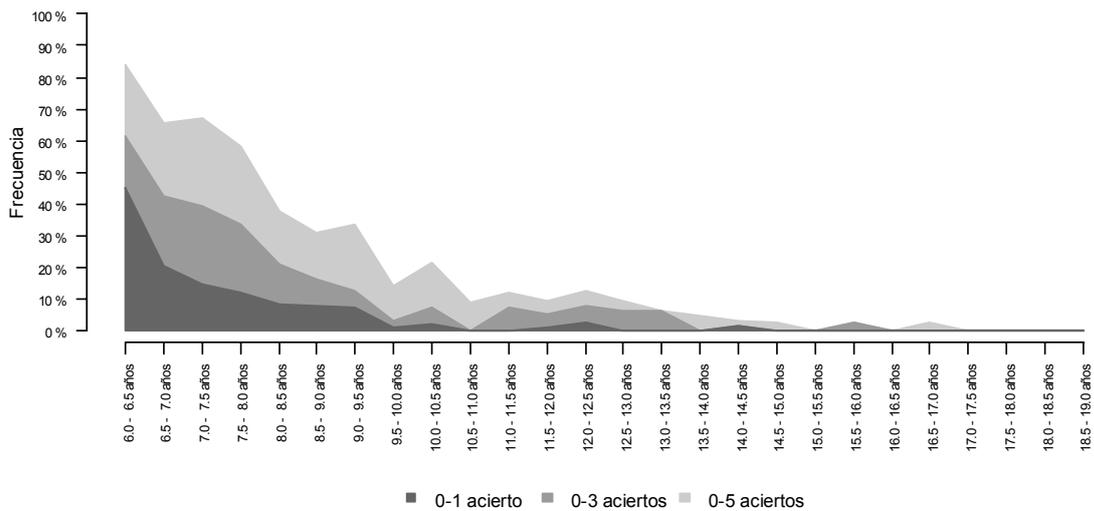
Consideramos el efecto suelo en una prueba cuando no se puede llegar a discriminar entre sujetos por debajo de una puntuación obtenida. En los siguientes gráficos se muestra los porcentajes de sujetos que han respondido correctamente a un número muy bajo de ítems según la edad en el momento de la aplicación, observamos como ese porcentaje disminuye al aumentar la edad de los sujetos.



Prueba Contextos - Efecto Suelo



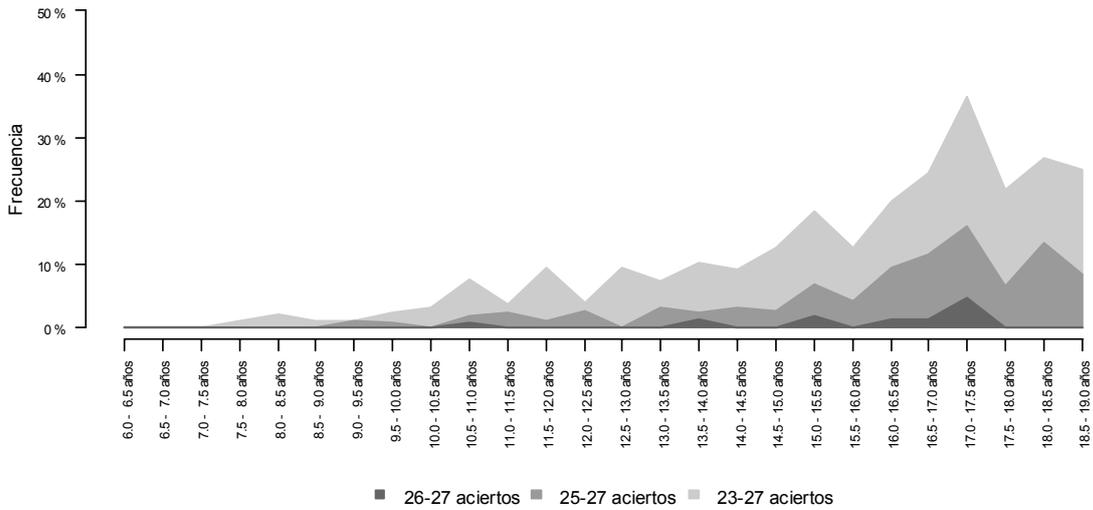
Prueba Diseños - Efecto Suelo



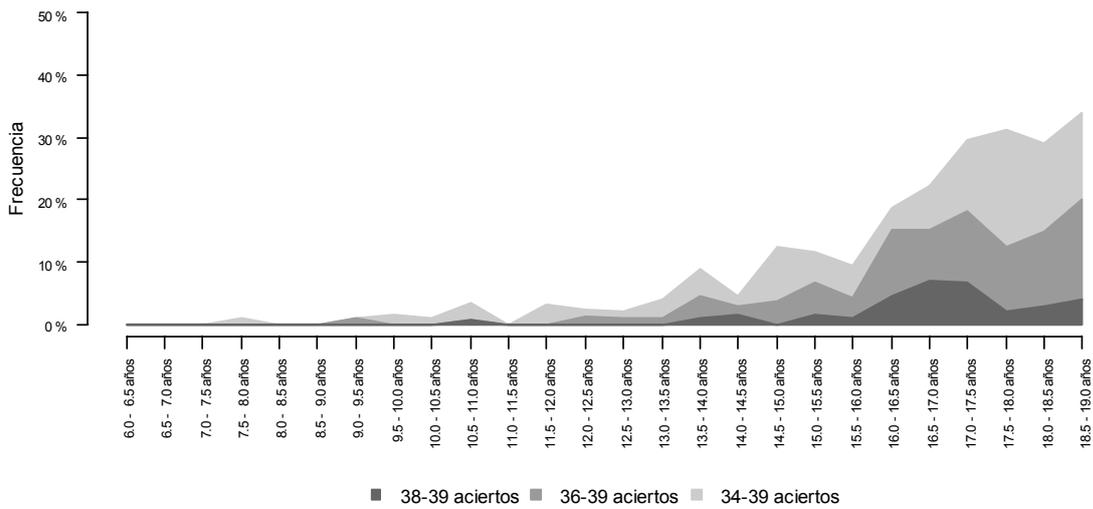
Consideramos el efecto techo en una prueba cuando no se puede llegar a discriminar entre sujetos por encima de una puntuación obtenida. En el siguiente gráfico se muestra qué porcentaje de sujetos, por grupo de edad, obtuvo una puntuación máxima en las pruebas de Semejanzas, Analogías y Contextos. Observamos cómo este efecto únicamente se produce en edades cercanas a los 18 años, edad límite de la prueba.

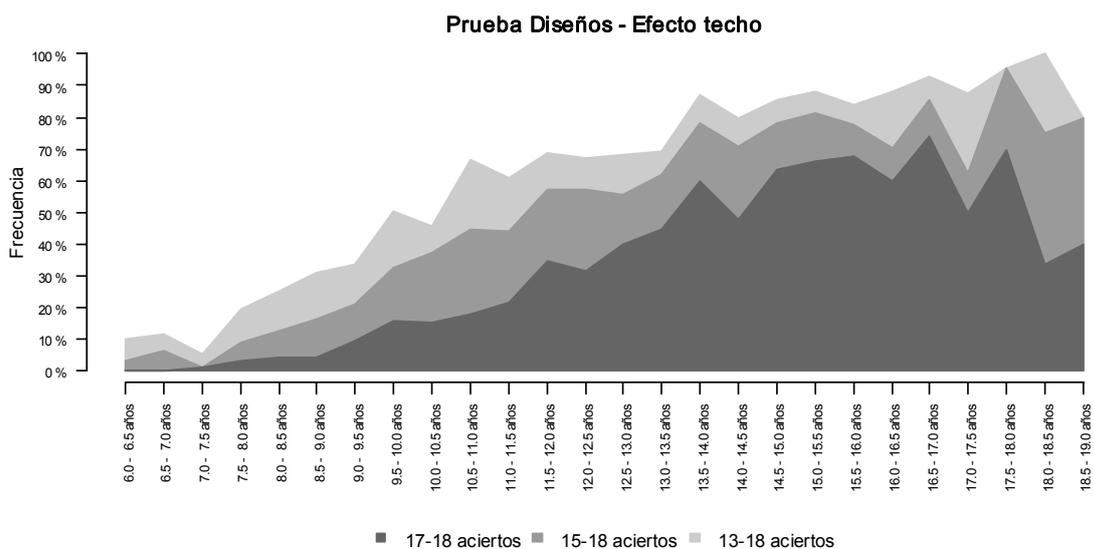
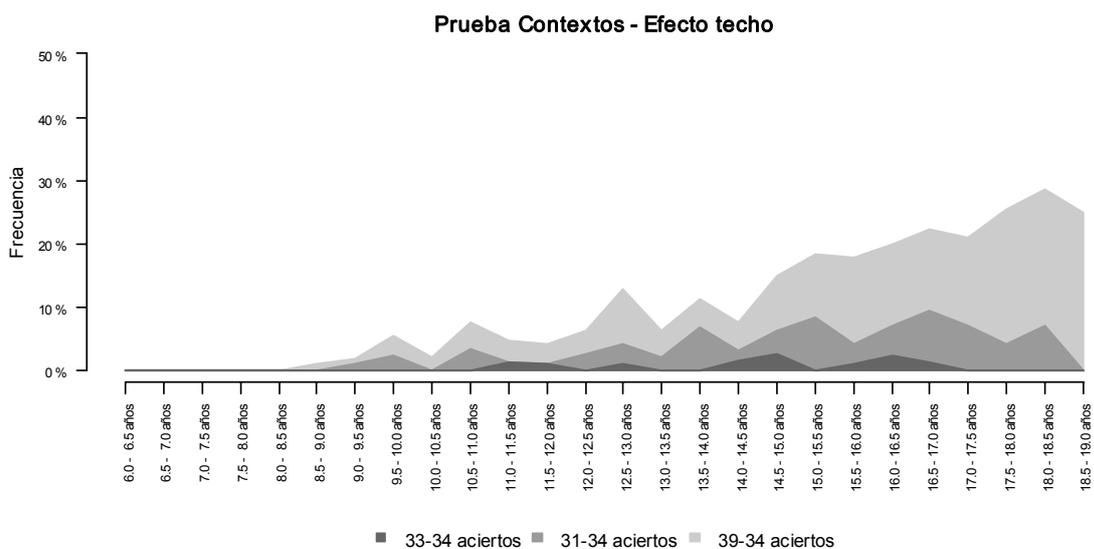


Prueba Semejanzas - Efecto techo



Prueba Analogías - Efecto techo





Se ha considerado que la prueba Diseños tiene un efecto techo que limita la edad de aplicación de la prueba a 12 años.