



Introducción a Modelos Psicométricos

Clase 15

Funcionamiento Diferencial de los Ítems

Iwin Leenen y Ramsés Vázquez-Lira

Facultad de Psicología, UNAM

Programa de Licenciatura y Posgrado en Psicología
Semestre 2019–1

Índice

- 1 Funcionamiento diferencial: Definición
- 2 El método de Mantel-Haenszel
- 3 Métodos basados en modelos TRI
- 4 Comparación de ambos métodos (Mantel-Haenszel y basados en la TRI)

Introducción

Funcionamiento diferencial de los ítems o del test

■ Punto de partida:

Las puntuaciones en muchas pruebas muestran diferencias significativas entre grupos sociales, entre hombres y mujeres, etc.

Por ejemplo:

- En México, los niños de escuelas indígenas tienen puntuaciones más bajas en PLANEA que los niños no indígenas.
- En EE.UU., los sustentantes afroamericanos tienen resultados más bajos en los exámenes de admisión para la universidad.
- En Europa, candidatos autóctonos que solicitan para un trabajo tienen resultados más bajos en las pruebas de selección.

■ Pregunta:

¿A qué se deben estas diferencias?

- ¿A diferencias reales en el rasgo (o los rasgos) que subyace(n) el test?
- ¿A diferencias no relacionadas con este(os) rasgo(s) latente(s)?
⇒ **Funcionamiento diferencial** del test o de (algunos) ítems.

Introducción

Funcionamiento diferencial de los ítems o del test

■ Punto de partida:

Las puntuaciones en muchas pruebas muestran diferencias significativas entre grupos sociales, entre hombres y mujeres, etc.

Por ejemplo:

- En México, los niños de escuelas indígenas tienen puntuaciones más bajas en PLANEA que los niños no indígenas.
- En EE.UU., los sustentantes afroamericanos tienen resultados más bajos en los exámenes de admisión para la universidad.
- En Europa, candidatos autóctonos que solicitan para un trabajo tienen resultados más bajos en las pruebas de selección.

■ Pregunta:

¿A qué se deben estas diferencias?

- ¿A diferencias reales en el rasgo (o los rasgos) que subyace(n) el test?
- ¿A diferencias no relacionadas con este(os) rasgo(s) latente(s)?
⇒ **Funcionamiento diferencial** del test o de (algunos) ítems.

Introducción

Funcionamiento diferencial de los ítems o del test

■ Punto de partida:

Las puntuaciones en muchas pruebas muestran diferencias significativas entre grupos sociales, entre hombres y mujeres, etc.

Por ejemplo:

- En México, los niños de escuelas indígenas tienen puntuaciones más bajas en PLANEA que los niños no indígenas.
- En EE.UU., los sustentantes afroamericanos tienen resultados más bajos en los exámenes de admisión para la universidad.
- En Europa, candidatos autóctonos que solicitan para un trabajo tienen resultados más bajos en las pruebas de selección.

■ Pregunta:

¿A qué se deben estas diferencias?

- ¿A diferencias reales en el rasgo (o los rasgos) que subyace(n) el test?
- ¿A diferencias no relacionadas con este(os) rasgo(s) latente(s)?
⇒ **Funcionamiento diferencial** del test o de (algunos) ítems.

Introducción

Funcionamiento diferencial de los ítems o del test

■ Punto de partida:

Las puntuaciones en muchas pruebas muestran diferencias significativas entre grupos sociales, entre hombres y mujeres, etc.

Por ejemplo:

- En México, los niños de escuelas indígenas tienen puntuaciones más bajas en PLANEA que los niños no indígenas.
- En EE.UU., los sustentantes afroamericanos tienen resultados más bajos en los exámenes de admisión para la universidad.
- En Europa, candidatos autóctonos que solicitan para un trabajo tienen resultados más bajos en las pruebas de selección.

■ Pregunta:

¿A qué se deben estas diferencias?

- ¿A diferencias reales en el rasgo (o los rasgos) que subyace(n) el test?
- ¿A diferencias no relacionadas con este(os) rasgo(s) latente(s)?
⇒ **Funcionamiento diferencial** del test o de (algunos) ítems.

Introducción

Funcionamiento diferencial de los ítems o del test

■ Punto de partida:

Las puntuaciones en muchas pruebas muestran diferencias significativas entre grupos sociales, entre hombres y mujeres, etc.

Por ejemplo:

- En México, los niños de escuelas indígenas tienen puntuaciones más bajas en PLANEA que los niños no indígenas.
- En EE.UU., los sustentantes afroamericanos tienen resultados más bajos en los exámenes de admisión para la universidad.
- En Europa, candidatos autóctonos que solicitan para un trabajo tienen resultados más bajos en las pruebas de selección.

■ Pregunta:

¿A qué se deben estas diferencias?

- ¿A diferencias reales en el rasgo (o los rasgos) que subyace(n) el test?
- ¿A diferencias no relacionadas con este(os) rasgo(s) latente(s)?
⇒ **Funcionamiento diferencial** del test o de (algunos) ítems.

Introducción

Funcionamiento diferencial de los ítems o del test

■ Punto de partida:

Las puntuaciones en muchas pruebas muestran diferencias significativas entre grupos sociales, entre hombres y mujeres, etc.

Por ejemplo:

- En México, los niños de escuelas indígenas tienen puntuaciones más bajas en PLANEA que los niños no indígenas.
- En EE.UU., los sustentantes afroamericanos tienen resultados más bajos en los exámenes de admisión para la universidad.
- En Europa, candidatos autóctonos que solicitan para un trabajo tienen resultados más bajos en las pruebas de selección.

■ Pregunta:

¿A qué se deben estas diferencias?

- ¿A diferencias reales en el rasgo (o los rasgos) que subyace(n) el test?
- ¿A diferencias no relacionadas con este(os) rasgo(s) latente(s)?
 ⇒ **Funcionamiento diferencial** del test o de (algunos) ítems.

Introducción

Funcionamiento diferencial de los ítems o del test

■ Punto de partida:

Las puntuaciones en muchas pruebas muestran diferencias significativas entre grupos sociales, entre hombres y mujeres, etc.

Por ejemplo:

- En México, los niños de escuelas indígenas tienen puntuaciones más bajas en PLANEA que los niños no indígenas.
- En EE.UU., los sustentantes afroamericanos tienen resultados más bajos en los exámenes de admisión para la universidad.
- En Europa, candidatos autóctonos que solicitan para un trabajo tienen resultados más bajos en las pruebas de selección.

■ Pregunta:

¿A qué se deben estas diferencias?

- ¿A diferencias reales en el rasgo (o los rasgos) que subyace(n) el test?
- ¿A diferencias no relacionadas con este(os) rasgo(s) latente(s)?

⇒ **Funcionamiento diferencial** del test o de (algunos) ítems.

Introducción

Funcionamiento diferencial de los ítems o del test

■ Punto de partida:

Las puntuaciones en muchas pruebas muestran diferencias significativas entre grupos sociales, entre hombres y mujeres, etc.

Por ejemplo:

- En México, los niños de escuelas indígenas tienen puntuaciones más bajas en PLANEA que los niños no indígenas.
- En EE.UU., los sustentantes afroamericanos tienen resultados más bajos en los exámenes de admisión para la universidad.
- En Europa, candidatos autóctonos que solicitan para un trabajo tienen resultados más bajos en las pruebas de selección.

■ Pregunta:

¿A qué se deben estas diferencias?

- ¿A diferencias reales en el rasgo (o los rasgos) que subyace(n) el test?
- ¿A diferencias no relacionadas con este(os) rasgo(s) latente(s)?
⇒ **Funcionamiento diferencial** del test o de (algunos) ítems.

Índice

- 1 **Funcionamiento diferencial: Definición**
- 2 El método de Mantel-Haenszel
- 3 Métodos basados en modelos TRI
- 4 Comparación de ambos métodos (Mantel-Haenszel y basados en la TRI)

Funcionamiento diferencial de los ítems y el test: Concepto

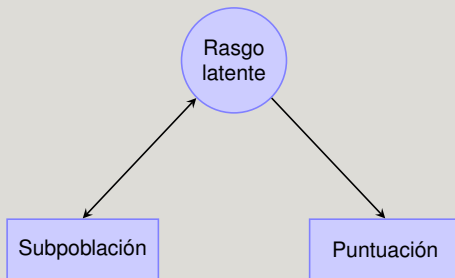
Ausencia de funcionamiento diferencial



- Se observa una asociación entre la pertenencia a una subpoblación y la puntuación observada en un ítem o test.

Funcionamiento diferencial de los ítems y el test: Concepto

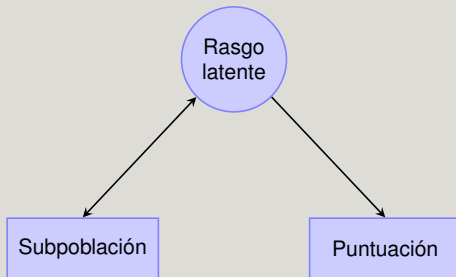
Ausencia de funcionamiento diferencial



- Se observa una asociación entre la pertenencia a una subpoblación y la puntuación observada en un ítem o test.
- Esta asociación se explica **por completo** por el rasgo latente que se pretende medir.

Funcionamiento diferencial de los ítems y el test: Concepto

Ausencia de funcionamiento diferencial

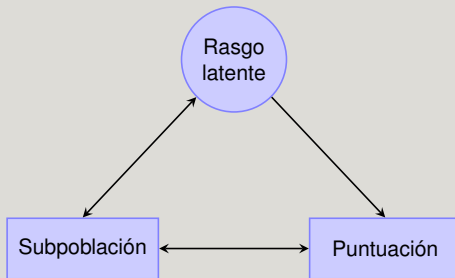


- Se observa una asociación entre la pertenencia a una subpoblación y la puntuación observada en un ítem o test.
- Esta asociación se explica **por completo** por el rasgo latente que se pretende medir.

⇒ Ítems o tests **insesgados** o **sin funcionamiento diferencial**.

Funcionamiento diferencial de los ítems y el test: Concepto

Tests o ítems con funcionamiento diferencial

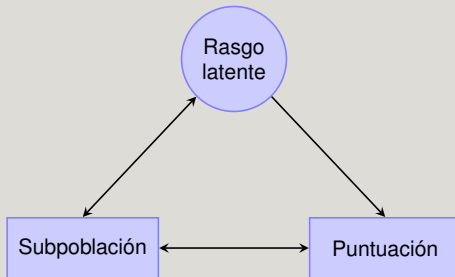


- Sin embargo, después de tomar en cuenta el rasgo latente, permanece una asociación **directa** entre la pertenencia a una subpoblación y la puntuación observada.

⇒ **Funcionamiento diferencial** de los ítems o del test.

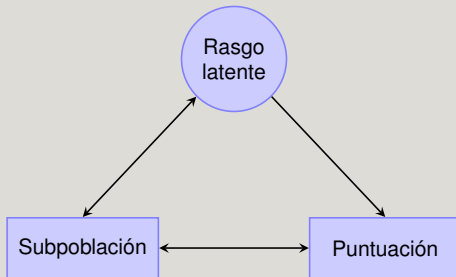
Funcionamiento diferencial de los ítems y el test: Concepto

Dos tipos de funcionamiento diferencial



Funcionamiento diferencial de los ítems y el test: Concepto

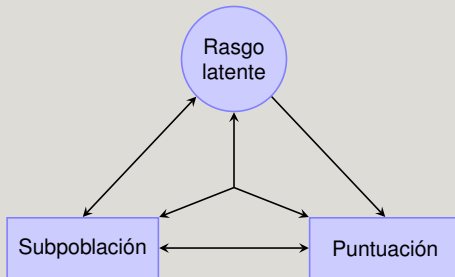
Dos tipos de funcionamiento diferencial



1. Las diferencias entre grupos son constantes para todos los niveles en el rasgo latente.
⇒ Funcionamiento diferencial **uniforme**.

Funcionamiento diferencial de los ítems y el test: Concepto

Dos tipos de funcionamiento diferencial



1. Las diferencias entre grupos son constantes para todos los niveles en el rasgo latente.
⇒ Funcionamiento diferencial **uniforme**.
2. Las diferencias entre grupos son diferentes entre los distintos niveles en el rasgo latente.
⇒ Funcionamiento diferencial **no uniforme**.

Funcionamiento diferencial de los ítems y el test: Definición

Definición

Un ítem (o un test) muestra **funcionamiento diferencial** si y solo si personas

- que pertenecen a **diferentes grupos**,
- pero que tienen **el mismo nivel en el rasgo latente** que se pretende medir, tienen una probabilidad diferente de obtener cierta puntuación en el ítem (o el test).

Funcionamiento diferencial de los ítems y el test

Funcionamiento diferencial del ítem vs funcionamiento diferencial del test

- **Conceptualmente**, es similar funcionamiento diferencial para la prueba en su totalidad o para uno o algunos ítems en el test.
 - DIF: Funcionamiento diferencial del ítem
 - DTF: Funcionamiento diferencial del test
- Sin embargo, respecto de los **métodos para detectar funcionamiento diferencial**, la diferencia entre DIF y DTF sí es importante:
 - Para detectar DTF, se requiere una medición adicional insesgada para encontrar las diferencias entre los grupos.
 - En caso de DIF (es decir, si solo un subconjunto de los ítems del test muestran funcionamiento diferencial), es posible detectarlo sin medición adicional.
- En estudios para detectar funcionamiento diferencial, es común utilizar la siguiente terminología:
 - **Grupo de referencia** (comúnmente el grupo mayoritario);
 - **Grupo focal** (comúnmente el grupo que posiblemente sufre una desventaja).

Funcionamiento diferencial de los ítems y el test

Funcionamiento diferencial del ítem vs funcionamiento diferencial del test

- **Conceptualmente**, es similar funcionamiento diferencial para la prueba en su totalidad o para uno o algunos ítems en el test.
 - DIF: Funcionamiento diferencial del ítem
 - DTF: Funcionamiento diferencial del test
- Sin embargo, respecto de los métodos para detectar funcionamiento diferencial, la diferencia entre DIF y DTF sí es importante:
 - Para detectar DTF, se requiere una medición adicional insesgada para encontrar las diferencias entre los grupos.
 - En caso de DIF (es decir, si solo un subconjunto de los ítems del test muestran funcionamiento diferencial), es posible detectarlo sin medición adicional.
- En estudios para detectar funcionamiento diferencial, es común utilizar la siguiente terminología:
 - Grupo de referencia (comúnmente el grupo mayoritario);
 - Grupo focal (comúnmente el grupo que posiblemente sufre una desventaja).

Funcionamiento diferencial de los ítems y el test

Funcionamiento diferencial del ítem vs funcionamiento diferencial del test

- **Conceptualmente**, es similar funcionamiento diferencial para la prueba en su totalidad o para uno o algunos ítems en el test.
 - DIF: Funcionamiento diferencial del ítem
 - DTF: Funcionamiento diferencial del test
- Sin embargo, respecto de los **métodos para detectar funcionamiento diferencial**, la diferencia entre DIF y DTF sí es importante:
 - Para detectar DTF, se requiere una medición adicional insesgada para encontrar las diferencias entre los grupos.
 - En caso de DIF (es decir, si solo un subconjunto de los ítems del test muestran funcionamiento diferencial), es posible detectarlo sin medición adicional.
- En estudios para detectar funcionamiento diferencial, es común utilizar la siguiente terminología:
 - **Grupo de referencia** (comúnmente el grupo mayoritario);
 - **Grupo focal** (comúnmente el grupo que posiblemente sufre una desventaja).

Funcionamiento diferencial de los ítems y el test

Funcionamiento diferencial del ítem vs funcionamiento diferencial del test

- **Conceptualmente**, es similar funcionamiento diferencial para la prueba en su totalidad o para uno o algunos ítems en el test.
 - DIF: Funcionamiento diferencial del ítem
 - DTF: Funcionamiento diferencial del test
- Sin embargo, respecto de los **métodos para detectar funcionamiento diferencial**, la diferencia entre DIF y DTF sí es importante:
 - Para detectar DTF, se requiere una medición adicional insesgada para encontrar las diferencias entre los grupos.
 - En caso de DIF (es decir, si solo un subconjunto de los ítems del test muestran funcionamiento diferencial), es posible detectarlo sin medición adicional.
- En estudios para detectar funcionamiento diferencial, es común utilizar la siguiente terminología:
 - **Grupo de referencia** (comúnmente el grupo mayoritario);
 - **Grupo focal** (comúnmente el grupo que posiblemente sufre una desventaja).

Funcionamiento diferencial de los ítems y el test

Funcionamiento diferencial del ítem vs funcionamiento diferencial del test

- **Conceptualmente**, es similar funcionamiento diferencial para la prueba en su totalidad o para uno o algunos ítems en el test.
 - DIF: Funcionamiento diferencial del ítem
 - DTF: Funcionamiento diferencial del test
- Sin embargo, respecto de los **métodos para detectar funcionamiento diferencial**, la diferencia entre DIF y DTF sí es importante:
 - Para detectar DTF, se requiere una medición adicional insesgada para encontrar las diferencias entre los grupos.
 - En caso de DIF (es decir, si solo un subconjunto de los ítems del test muestran funcionamiento diferencial), es posible detectarlo sin medición adicional.
- En estudios para detectar funcionamiento diferencial, es común utilizar la siguiente terminología:
 - **Grupo de referencia** (comúnmente el grupo mayoritario);
 - **Grupo focal** (comúnmente el grupo que posiblemente sufre una desventaja).

Índice

- 1 Funcionamiento diferencial: Definición
- 2 El método de Mantel-Haenszel**
- 3 Métodos basados en modelos TRI
- 4 Comparación de ambos métodos (Mantel-Haenszel y basados en la TRI)

El método de Mantel-Haenszel

Mantel-Haenszel: Procedimiento

- El procedimiento de Mantel-Haenszel utiliza la puntuación total en el test como proxy del nivel en el rasgo latente.
- Aquí se considera el caso para detectar funcionamiento diferencial de ítems dicotómicos.
- Para detectar DIF en cierto ítem, se construye la siguiente tabla de 2×2 para cada grupo de nivel q :

	Puntuación en el ítem		Total
	1 (acierto)	0 (fallo)	
Grupo de referencia	a_q	b_q	n_{Rq}
Grupo focal	c_q	d_q	n_{Fq}
Total	n_{1q}	n_{0q}	n_q

El método de Mantel-Haenszel

Mantel-Haenszel: Procedimiento

- El procedimiento de Mantel-Haenszel utiliza la puntuación total en el test como proxy del nivel en el rasgo latente.
- Aquí se considera el caso para detectar funcionamiento diferencial de ítems dicotómicos.
- Para detectar DIF en cierto ítem, se construye la siguiente tabla de 2×2 para cada grupo de nivel q :

	Puntuación en el ítem		Total
	1 (acierto)	0 (fallo)	
Grupo de referencia	a_q	b_q	n_{Rq}
Grupo focal	c_q	d_q	n_{Fq}
Total	n_{1q}	n_{0q}	n_q

El método de Mantel-Haenszel

Mantel-Haenszel: Procedimiento

- El procedimiento de Mantel-Haenszel utiliza la puntuación total en el test como proxy del nivel en el rasgo latente.
- Aquí se considera el caso para detectar funcionamiento diferencial de ítems dicotómicos.
- Para detectar DIF en cierto ítem, se construye la siguiente tabla de 2×2 para cada grupo de nivel q :

	Puntuación en el ítem		Total
	1 (acierto)	0 (fallo)	
Grupo de referencia	a_q	b_q	n_{Rq}
Grupo focal	c_q	d_q	n_{Fq}
Total	n_{1q}	n_{0q}	n_q

El método de Mantel-Haenszel

Mantel-Haenszel: Procedimiento

	Puntuación en el ítem		Total
	1 (acierto)	0 (fallo)	
Grupo de referencia	a_q	b_q	n_{Rq}
Grupo focal	c_q	d_q	n_{Fq}
Total	n_{1q}	n_{0q}	n_q

- Se considera la *odds ratio* (razón de momios) α_q para cada nivel q :

$$\alpha_q = \frac{\text{momio de acertar en el grupo de referencia}}{\text{momio de acertar en el grupo focal}} = \frac{\text{momio}_{Rq}}{\text{momio}_{Fq}}$$

El momio de acertar en cualquier de los grupos se define por:

$$\frac{\text{probabilidad de acertar}}{\text{probabilidad de fallar}} \quad \text{para este grupo.}$$

El método de Mantel-Haenszel

Mantel-Haenszel: Procedimiento

	Puntuación en el ítem		Total
	1 (acierto)	0 (fallo)	
Grupo de referencia	a_q	b_q	n_{Rq}
Grupo focal	c_q	d_q	n_{Fq}
Total	n_{1q}	n_{0q}	n_q

- Se considera la *odds ratio* (razón de momios) α_q para cada nivel q :

$$\alpha_q = \frac{\text{momio de acertar en el grupo de referencia}}{\text{momio de acertar en el grupo focal}} = \frac{\text{momio}_{Rq}}{\text{momio}_{Fq}}$$

Por lo tanto, se estiman por:

$$\hat{\text{momio}}_{Rq} = \frac{a_q / n_{Rq}}{b_q / n_{Rq}} = \frac{a_q}{b_q} \quad \text{y} \quad \hat{\text{momio}}_{Fq} = \frac{c_q / n_{Fq}}{d_q / n_{Fq}} = \frac{c_q}{d_q}$$

El método de Mantel-Haenszel

Mantel-Haenszel: Procedimiento

	Puntuación en el ítem		Total
	1 (acierto)	0 (fallo)	
Grupo de referencia	a_q	b_q	n_{Rq}
Grupo focal	c_q	d_q	n_{Fq}
Total	n_{1q}	n_{0q}	n_q

- Se considera la *odds ratio* (razón de momios) α_q para cada nivel q :

$$\alpha_q = \frac{\text{momio de acertar en el grupo de referencia}}{\text{momio de acertar en el grupo focal}} = \frac{\text{momio}_{Rq}}{\text{momio}_{Fq}}$$

Y se estima α_q por:

$$\hat{\alpha}_q = \frac{a_q / b_q}{c_q / d_q} = \frac{a_q d_q}{b_q c_q}$$

El método de Mantel-Haenszel

Mantel-Haenszel: Procedimiento

- Se extiende la lógica anterior para obtener una razón de momios combinada de los Q grupos de nivel:

$$\hat{\alpha}_{\text{MH}} = \frac{\sum_{q=1}^Q \frac{a_q d_q}{n_q}}{\sum_{q=1}^Q \frac{b_q c_q}{n_q}}$$

- Interpretación:

$\alpha_{\text{MH}} > 1 \implies$ La probabilidad de acertar es más alta en el grupo de referencia

$\alpha_{\text{MH}} = 1 \implies$ La probabilidad de acertar es igual en ambos grupos

$\alpha_{\text{MH}} < 1 \implies$ La probabilidad de acertar es más alta en el grupo focal

- Se puede realizar una prueba de la hipótesis nula:

$$\mathcal{H}_0 : \alpha_{\text{MH}} = 1 \quad (\text{No hay DIF}),$$

con base en:

$$\log \hat{\alpha}_{\text{MH}} \stackrel{\text{aprox}}{\sim} N\left(0, \sigma_{\log \hat{\alpha}_{\text{MH}}}^2\right).$$

El método de Mantel-Haenszel

Mantel-Haenszel: Procedimiento

- Se extiende la lógica anterior para obtener una razón de momios combinada de los Q grupos de nivel:

$$\hat{\alpha}_{MH} = \frac{\sum_{q=1}^Q \frac{a_q d_q}{n_q}}{\sum_{q=1}^Q \frac{b_q c_q}{n_q}}$$

- Interpretación:

$\alpha_{MH} > 1 \implies$ La probabilidad de acertar es más alta en el grupo de referencia

$\alpha_{MH} = 1 \implies$ La probabilidad de acertar es igual en ambos grupos

$\alpha_{MH} < 1 \implies$ La probabilidad de acertar es más alta en el grupo focal

- Se puede realizar una prueba de la hipótesis nula:

$$\mathcal{H}_0 : \alpha_{MH} = 1 \quad (\text{No hay DIF}),$$

con base en:

$$\log \hat{\alpha}_{MH} \stackrel{\text{aprox}}{\sim} N\left(0, \sigma_{\log \hat{\alpha}_{MH}}^2\right).$$

El método de Mantel-Haenszel

Mantel-Haenszel: Procedimiento

- Se extiende la lógica anterior para obtener una razón de momios combinada de los Q grupos de nivel:

$$\hat{\alpha}_{\text{MH}} = \frac{\sum_{q=1}^Q \frac{a_q d_q}{n_q}}{\sum_{q=1}^Q \frac{b_q c_q}{n_q}}$$

- Interpretación:

$\alpha_{\text{MH}} > 1 \implies$ La probabilidad de acertar es más alta en el grupo de referencia

$\alpha_{\text{MH}} = 1 \implies$ La probabilidad de acertar es igual en ambos grupos

$\alpha_{\text{MH}} < 1 \implies$ La probabilidad de acertar es más alta en el grupo focal

- Se puede realizar una prueba de la hipótesis nula:

$$\mathcal{H}_0 : \alpha_{\text{MH}} = 1 \quad (\text{No hay DIF}),$$

con base en:

$$\log \hat{\alpha}_{\text{MH}} \stackrel{\text{aprox}}{\sim} N\left(0, \sigma_{\log \hat{\alpha}_{\text{MH}}}^2\right).$$

El método de Mantel-Haenszel

Mantel-Haenszel: Consideraciones finales

- La presencia de ítems con DIF compromete la puntuación total en el test como proxy del nivel en el rasgo latente.
- Por lo tanto, encontrar los ítems con DIF es el resultado de un proceso iterativo:
 - Se *inicia* el proceso con un análisis utilizando la puntuación total incluyendo *todos los ítems*.
 - A continuación, se realiza el análisis utilizando la puntuación total incluyendo solo los ítems del paso previo *que no exhibieron DIF*
Es posible que se detecten nuevos ítems con DIF
y/o ítems previamente detectado con DIF ya no exhiben DIF
 - Se continua el proceso hasta que:
 - ninguno de los ítems incluidos en el cálculo de la puntuación total exhiban DIF;
 - todos los demás ítems sí exhiban DIF.

El método de Mantel-Haenszel

Mantel-Haenszel: Consideraciones finales

- La presencia de ítems con DIF compromete la puntuación total en el test como proxy del nivel en el rasgo latente.
- Por lo tanto, encontrar los ítems con DIF es el resultado de un proceso iterativo:
 - Se *inicia* el proceso con un análisis utilizando la puntuación total incluyendo *todos los ítems*.
 - A continuación, se realiza el análisis utilizando la puntuación total incluyendo solo los ítems del paso previo *que no exhibieron DIF*
Es posible que se detecten nuevos ítems con DIF
y/o ítems previamente detectado con DIF ya no exhiben DIF
 - Se continua el proceso hasta que:
 - ninguno de los ítems incluidos en el cálculo de la puntuación total exhiban DIF;
 - todos los demás ítems sí exhiban DIF.

El método de Mantel-Haenszel

Mantel-Haenszel: Consideraciones finales

- La presencia de ítems con DIF compromete la puntuación total en el test como proxy del nivel en el rasgo latente.
- Por lo tanto, encontrar los ítems con DIF es el resultado de un proceso iterativo:
 - Se **inicia** el proceso con un análisis utilizando la puntuación total incluyendo *todos los ítems*.
 - A continuación, se realiza el análisis utilizando la puntuación total incluyendo solo los ítems del paso previo **que no exhibieron DIF**
Es posible que se detecten nuevos ítems con DIF
y/o ítems previamente detectado con DIF ya no exhiben DIF
 - Se continua el proceso hasta que:
 - ninguno de los ítems incluidos en el cálculo de la puntuación total exhiban DIF;
 - todos los demás ítems sí exhiban DIF.

El método de Mantel-Haenszel

Mantel-Haenszel: Consideraciones finales

- La presencia de ítems con DIF compromete la puntuación total en el test como proxy del nivel en el rasgo latente.
- Por lo tanto, encontrar los ítems con DIF es el resultado de un proceso iterativo:
 - Se **inicia** el proceso con un análisis utilizando la puntuación total incluyendo *todos los ítems*.
 - A continuación, se realiza el análisis utilizando la puntuación total incluyendo solo los ítems del paso previo **que no exhibieron DIF**
Es posible que se detecten nuevos ítems con DIF
y/o ítems previamente detectado con DIF ya no exhiben DIF
 - Se continua el proceso hasta que:
 - ninguno de los ítems incluidos en el cálculo de la puntuación total exhiban DIF;
 - todos los demás ítems sí exhiban DIF.

El método de Mantel-Haenszel

Mantel-Haenszel: Consideraciones finales

- La presencia de ítems con DIF compromete la puntuación total en el test como proxy del nivel en el rasgo latente.
- Por lo tanto, encontrar los ítems con DIF es el resultado de un proceso iterativo:
 - Se **inicia** el proceso con un análisis utilizando la puntuación total incluyendo *todos los ítems*.
 - A continuación, se realiza el análisis utilizando la puntuación total incluyendo solo los ítems del paso previo **que no exhibieron DIF**
Es posible que se detecten nuevos ítems con DIF
y/o ítems previamente detectado con DIF ya no exhiben DIF
 - Se continua el proceso hasta que:
 - ninguno de los ítems incluidos en el cálculo de la puntuación total exhiban DIF;
 - todos los demás ítems sí exhiban DIF.

Índice

- 1 Funcionamiento diferencial: Definición
- 2 El método de Mantel-Haenszel
- 3 Métodos basados en modelos TRI**
- 4 Comparación de ambos métodos (Mantel-Haenszel y basados en la TRI)

Funcionamiento diferencial del ítem en modelos TRI

DIF en la TRI: Ideas generales

- En la TRI, la probabilidad de que una persona p , al contestar el ítem i , dé una respuesta en la categoría j es una función de los parámetros de los ítems y de las personas.
- En principio, dos personas cuyos parámetros son iguales tendrán la misma probabilidad de contestar en las distintas categorías de cualquier ítem **independientemente de la subpoblación a la que pertenecen**.
- El el marco de la TRI, se considera que un ítem exhibe DIF si
 - dos individuos que pertenecen a diferentes subpoblaciones tienen probabilidades distintas para responder en cierta categoría de respuesta del ítem **aunque tienen el mismo nivel en el rasgo latente**.
 - los parámetros del ítem varían entre diferentes subpoblaciones de individuos (suponiendo que los parámetros se han calibrado sobre la misma dimensión).

Funcionamiento diferencial del ítem en modelos TRI

DIF en la TRI: Ideas generales

- En la TRI, la probabilidad de que una persona p , al contestar el ítem i , dé una respuesta en la categoría j es una función de los parámetros de los ítems y de las personas.
- En principio, dos personas cuyos parámetros son iguales tendrán la misma probabilidad de contestar en las distintas categorías de cualquier ítem **independientemente de la subpoblación a la que pertenecen**.
- El el marco de la TRI, se considera que un ítem exhibe DIF si
 - dos individuos que pertenecen a diferentes subpoblaciones tienen probabilidades distintas para responder en cierta categoría de respuesta del ítem **aunque tienen el mismo nivel en el rasgo latente**.
 - los parámetros del ítem varían entre diferentes subpoblaciones de individuos (suponiendo que los parámetros se han calibrado sobre la misma dimensión).

Funcionamiento diferencial del ítem en modelos TRI

DIF en la TRI: Ideas generales

- En la TRI, la probabilidad de que una persona p , al contestar el ítem i , dé una respuesta en la categoría j es una función de los parámetros de los ítems y de las personas.
- En principio, dos personas cuyos parámetros son iguales tendrán la misma probabilidad de contestar en las distintas categorías de cualquier ítem **independientemente de la subpoblación a la que pertenecen**.
- El el marco de la TRI, se considera que un ítem exhibe DIF si
 - dos individuos que pertenecen a diferentes subpoblaciones tienen probabilidades distintas para responder en cierta categoría de respuesta del ítem **aunque tienen el mismo nivel en el rasgo latente**.
 - los parámetros del ítem varían entre diferentes subpoblaciones de individuos (suponiendo que los parámetros se han calibrado sobre la misma dimensión).

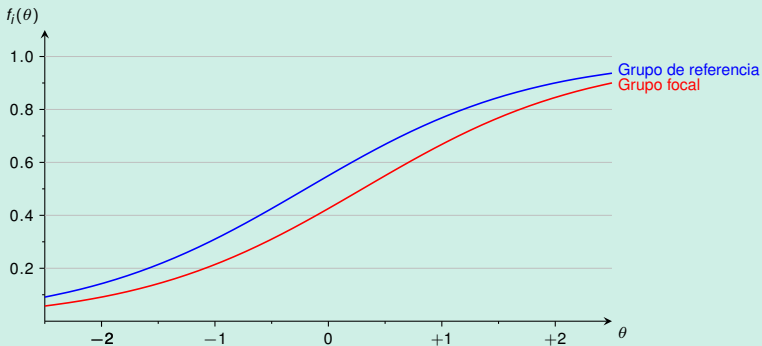
Funcionamiento diferencial del ítem en modelos TRI

DIF en la TRI: Ideas generales (continuación)

- La definición de DIF en el marco de la TRI es la misma, **independientemente del modelo concreto que se utiliza**:
 - Para los modelos para ítems dicotómicos de Rasch, 2PL, 3PL, ...
 - Para modelos para ítems polotómicos
 - Para modelos que suponen otra forma para la curva característica del ítem
 - Para modelos multidimensionales
 - ...
- Conceptualmente, es más fácil explicar e ilustrar DIF para **ítems dicotómicos en modelos unidimensionales**.

Funcionamiento diferencial del ítem en modelos TRI

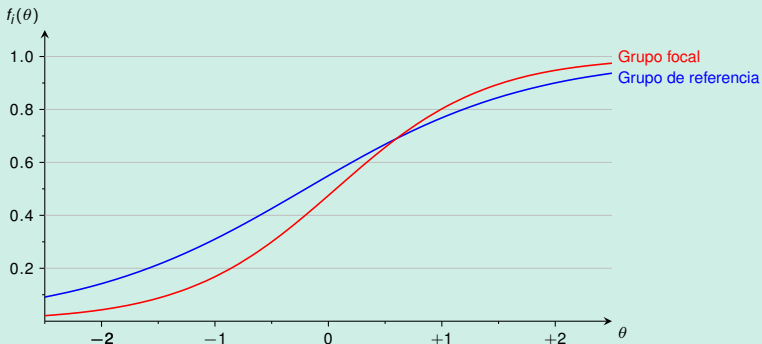
DIF uniforme en modelos unidimensionales para ítems dicotómicos la TRI



Ejemplo de la curva característica de un ítem con **DIF uniforme**.

Funcionamiento diferencial del ítem en modelos TRI

DIF no uniforme en modelos unidimensionales para ítems dicotómicos la TRI



Ejemplo de la curva característica de un ítem con **DIF no uniforme**.

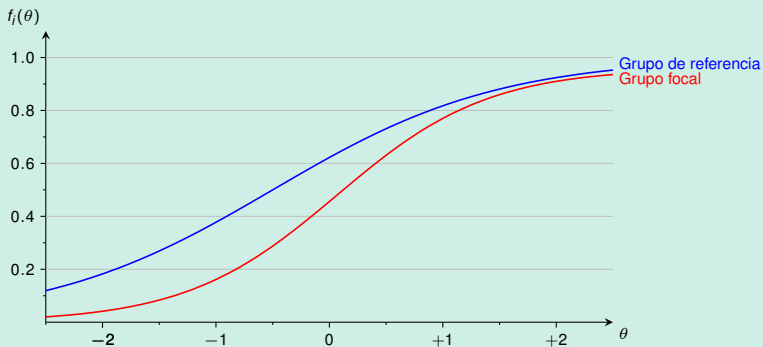
Funcionamiento diferencial del ítem en modelos TRI

DIF en modelos unidimensionales para ítems dicotómicos la TRI

- En líneas generales, la diferencia entre DIF uniforme y DIF no uniforme se relaciona con que las curvas características se cruzan o no.
Sin embargo,... véase la figura de la siguiente diapositiva.
- En los ejemplos anteriores, es relativamente fácil modelar el DIF a través de los parámetros de dificultad y discriminación del ítem.
- En la práctica, el patrón de DIF puede ser mucho más difícil para modelar.

Funcionamiento diferencial del ítem en modelos TRI

DIF no uniforme en modelos unidimensionales para ítems dicotómicos la TRI



Otro ejemplo de la curva característica de un ítem con **DIF no uniforme**.

Funcionamiento diferencial del ítem en modelos TRI

DIF en modelos unidimensionales para ítems dicotómicos la TRI

- En líneas generales, la diferencia entre DIF uniforme y DIF no uniforme se relaciona con que las curvas características se cruzan o no.
Sin embargo,... véase la figura de la siguiente diapositiva.
- En los ejemplos anteriores, es relativamente fácil modelar el DIF a través de los parámetros de dificultad y discriminación del ítem.
- En la práctica, el patrón de DIF puede ser mucho más difícil para modelar.

Funcionamiento diferencial del ítem en modelos TRI

DIF en modelos unidimensionales para ítems dicotómicos la TRI

- En líneas generales, la diferencia entre DIF uniforme y DIF no uniforme se relaciona con que las curvas características se cruzan o no.
Sin embargo,... véase la figura de la siguiente diapositiva.
- En los ejemplos anteriores, es relativamente fácil modelar el DIF a través de los parámetros de dificultad y discriminación del ítem.
- En la práctica, el patrón de DIF puede ser mucho más difícil para modelar.

Funcionamiento diferencial del ítem en modelos TRI

Investigar funcionamiento diferencial del ítem con métodos de la TRI

El procedimiento para detectar funcionamiento diferencial de ítems en el marco de la TRI generalmente consiste en dos pasos:

1. Encontrar un modelo TRI con un ajuste satisfactorio a:
 - en primera instancia, los datos del grupo de referencia
 - y, si es posible, también a los datos del grupo focal.
2. Para cada ítem, analizar las diferencias entre el grupo de referencia y el grupo focal respecto de las probabilidades de acertar el ítem dentro de cada uno de distintos grupos de personas con el mismo nivel en el rasgo latente.

El análisis típicamente incluye:

- una prueba estadística de la hipótesis nula de que “no hay diferencia entre cierto(s) parámetro(s) del ítem” entre el grupo de referencia y el grupo focal.

Por ejemplo, $\mathcal{H}_0 : \beta_{IR} = \beta_{IF}$.

- un análisis gráfico, comparando las curvas características en ambos grupos.

Funcionamiento diferencial del ítem en modelos TRI

Investigar funcionamiento diferencial del ítem con métodos de la TRI

El procedimiento para detectar funcionamiento diferencial de ítems en el marco de la TRI generalmente consiste en dos pasos:

1. Encontrar un **modelo TRI con un ajuste satisfactorio** a:
 - en primera instancia, los datos del grupo de referencia
 - y, si es posible, también a los datos del grupo focal.
2. Para cada ítem, **analizar las diferencias** entre el grupo de referencia y el grupo focal respecto de las probabilidades de acertar el ítem dentro de cada uno de distintos grupos de personas con el mismo nivel en el rasgo latente.

El análisis típicamente incluye:

- una prueba estadística de la hipótesis nula de que “no hay diferencia entre cierto(s) parámetro(s) del ítem” entre el grupo de referencia y el grupo focal.

Por ejemplo, $\mathcal{H}_0 : \beta_{IR} = \beta_{IF}$.

- un análisis gráfico, comparando las curvas características en ambos grupos.

Funcionamiento diferencial del ítem en modelos TRI

Investigar funcionamiento diferencial del ítem con métodos de la TRI

El procedimiento para detectar funcionamiento diferencial de ítems en el marco de la TRI generalmente consiste en dos pasos:

1. Encontrar un **modelo TRI con un ajuste satisfactorio** a:
 - en primera instancia, los datos del grupo de referencia
 - y, si es posible, también a los datos del grupo focal.
2. Para cada ítem, **analizar las diferencias** entre el grupo de referencia y el grupo focal respecto de las probabilidades de acertar el ítem dentro de cada uno de distintos grupos de personas con el mismo nivel en el rasgo latente.

El análisis típicamente incluye:

- una prueba estadística de la hipótesis nula de que “no hay diferencia entre cierto(s) parámetro(s) del ítem” entre el grupo de referencia y el grupo focal.

Por ejemplo, $\mathcal{H}_0 : \beta_{IR} = \beta_{IF}$.

- un análisis gráfico, comparando las curvas características en ambos grupos.

Funcionamiento diferencial del ítem en modelos TRI

Investigar funcionamiento diferencial del ítem con métodos de la TRI

El procedimiento para detectar funcionamiento diferencial de ítems en el marco de la TRI generalmente consiste en dos pasos:

1. Encontrar un **modelo TRI con un ajuste satisfactorio** a:
 - en primera instancia, los datos del grupo de referencia
 - y, si es posible, también a los datos del grupo focal.
2. Para cada ítem, **analizar las diferencias** entre el grupo de referencia y el grupo focal respecto de las probabilidades de acertar el ítem dentro de cada uno de distintos grupos de personas con el mismo nivel en el rasgo latente.

El análisis típicamente incluye:

- una prueba estadística de la hipótesis nula de que “no hay diferencia entre cierto(s) parámetro(s) del ítem” entre el grupo de referencia y el grupo focal.

Por ejemplo, $\mathcal{H}_0 : \beta_{iR} = \beta_{iF}$.

- un análisis gráfico, comparando las curvas características en ambos grupos.

Funcionamiento diferencial del ítem en modelos TRI

Investigar funcionamiento diferencial del ítem con métodos de la TRI

El procedimiento para detectar funcionamiento diferencial de ítems en el marco de la TRI generalmente consiste en dos pasos:

1. Encontrar un **modelo TRI con un ajuste satisfactorio** a:
 - en primera instancia, los datos del grupo de referencia
 - y, si es posible, también a los datos del grupo focal.
2. Para cada ítem, **analizar las diferencias** entre el grupo de referencia y el grupo focal respecto de las probabilidades de acertar el ítem dentro de cada uno de distintos grupos de personas con el mismo nivel en el rasgo latente.

El análisis típicamente incluye:

- una prueba estadística de la hipótesis nula de que “no hay diferencia entre cierto(s) parámetro(s) del ítem” entre el grupo de referencia y el grupo focal.

Por ejemplo, $\mathcal{H}_0 : \beta_{iR} = \beta_{iF}$.

- un análisis gráfico, comparando las curvas características en ambos grupos.

Funcionamiento diferencial del ítem en modelos TRI

Investigar funcionamiento diferencial del ítem con métodos de la TRI (continuación)

En caso de que se detecten ítems con DIF, se suelen hacer dos pasos más:

3. Modelar las respuestas en los ítems con DIF con base en la muestra de personas de la población focal.
4. Evaluar las consecuencias de la presencia de los ítems con DIF en los resultados; particularmente, en la distribución de las θ s de las personas en la población focal.

Funcionamiento diferencial del ítem en modelos TRI

Investigar funcionamiento diferencial del ítem con métodos de la TRI (continuación)

En caso de que se detecten ítems con DIF, se suelen hacer dos pasos más:

3. Modelar las respuestas en los ítems con DIF con base en la muestra de personas de la población focal.
4. Evaluar las consecuencias de la presencia de los ítems con DIF en los resultados; particularmente, en la distribución de las θ s de las personas en la población focal.

Funcionamiento diferencial del ítem en modelos TRI

Investigar funcionamiento diferencial del ítem con métodos de la TRI (continuación)

En caso de que se detecten ítems con DIF, se suelen hacer dos pasos más:

3. Modelar las respuestas en los ítems con DIF con base en la muestra de personas de la población focal.
4. Evaluar las consecuencias de la presencia de los ítems con DIF en los resultados; particularmente, en la distribución de las θ s de las personas en la población focal.

Índice

- 1 Funcionamiento diferencial: Definición
- 2 El método de Mantel-Haenszel
- 3 Métodos basados en modelos TRI
- 4 Comparación de ambos métodos (Mantel-Haenszel y basados en la TRI)

Mantel-Haenszel vs la TRI para detectar DIF

- Supongamos que, en el marco de la TRI, se ha ajustado el modelo de Rasch y se utiliza este modelo para detectar DIF.

En este caso, el principio para dividir la muestra total en “grupos de nivel” es el mismo: la puntuación total en el test.

- Al comparar ambos métodos, se destacan las siguientes diferencias:
 - A un nivel teórico-conceptual, con el método de Mantel-Haenszel *no se ha validado la puntuación total* como medida del nivel en el rasgo latente.
 - El método de Mantel-Haenszel *no permite detectar ciertas formas de DIF* (especialmente DIF no uniforme).
 - En el modelo de Rasch, especialmente si se utiliza CML para estimar los parámetros de los ítems, los resultados para evaluar DIF en diferentes ítems son (más) *interdependientes*.
Por ejemplo, la presencia de ítems que perjudican el grupo focal conlleva, como artefacto, a que en otros ítems se presenta una ventaja para este grupo.

- → Recomendación: Utilizar ambos enfoques como métodos complementarios.

Mantel-Haenszel vs la TRI para detectar DIF

- Supongamos que, en el marco de la TRI, se ha ajustado el modelo de Rasch y se utiliza este modelo para detectar DIF.

En este caso, el principio para dividir la muestra total en “grupos de nivel” es el mismo: la puntuación total en el test.

- Al comparar ambos métodos, se destacan las siguientes diferencias:
 - A un nivel teórico-conceptual, con el método de Mantel-Haenszel *no se ha validado la puntuación total* como medida del nivel en el rasgo latente.
 - El método de Mantel-Haenszel *no permite detectar ciertas formas de DIF* (especialmente DIF no uniforme).
 - En el *modelo de Rasch*, especialmente si se utiliza CML para estimar los parámetros de los ítems, los resultados para evaluar DIF en diferentes ítems son (más) *interdependientes*.
Por ejemplo, la presencia de ítems que perjudican el grupo focal conlleva, como artefacto, a que en otros ítems se presenta una ventaja para este grupo.

- → Recomendación: Utilizar ambos enfoques como métodos complementarios.

Mantel-Haenszel vs la TRI para detectar DIF

- Supongamos que, en el marco de la TRI, se ha ajustado el modelo de Rasch y se utiliza este modelo para detectar DIF.

En este caso, el principio para dividir la muestra total en “grupos de nivel” es el mismo: la puntuación total en el test.

- Al comparar ambos métodos, se destacan las siguientes diferencias:
 - A un nivel teórico-conceptual, con el método de Mantel-Haenszel **no se ha validado la puntuación total** como medida del nivel en el rasgo latente.
 - El método de Mantel-Haenszel **no permite detectar ciertas formas de DIF** (especialmente DIF no uniforme).
 - En el modelo de Rasch, especialmente si se utiliza CML para estimar los parámetros de los ítems, los resultados para evaluar DIF en diferentes ítems son (más) **interdependientes**.
Por ejemplo, la presencia de ítems que perjudican el grupo focal conlleva, como artefacto, a que en otros ítems se presenta una ventaja para este grupo.

- → Recomendación: Utilizar ambos enfoques como métodos complementarios.

Mantel-Haenszel vs la TRI para detectar DIF

- Supongamos que, en el marco de la TRI, se ha ajustado el modelo de Rasch y se utiliza este modelo para detectar DIF.

En este caso, el principio para dividir la muestra total en “grupos de nivel” es el mismo: la puntuación total en el test.

- Al comparar ambos métodos, se destacan las siguientes diferencias:
 - A un nivel teórico-conceptual, con el método de Mantel-Haenszel **no se ha validado la puntuación total** como medida del nivel en el rasgo latente.
 - El método de **Mantel-Haenszel no permite detectar ciertas formas de DIF** (especialmente DIF no uniforme).
 - **En el modelo de Rasch**, especialmente si se utiliza CML para estimar los parámetros de los ítems, los resultados para evaluar DIF en diferentes ítems son (más) **interdependientes**.
Por ejemplo, la presencia de ítems que perjudican el grupo focal conlleva, como artefacto, a que en otros ítems se presenta una ventaja para este grupo.

- → Recomendación: Utilizar ambos enfoques como métodos complementarios.

Mantel-Haenszel vs la TRI para detectar DIF

- Supongamos que, en el marco de la TRI, se ha ajustado el modelo de Rasch y se utiliza este modelo para detectar DIF.

En este caso, el principio para dividir la muestra total en “grupos de nivel” es el mismo: la puntuación total en el test.

- Al comparar ambos métodos, se destacan las siguientes diferencias:
 - A un nivel teórico-conceptual, con el método de Mantel-Haenszel **no se ha validado la puntuación total** como medida del nivel en el rasgo latente.
 - El método de **Mantel-Haenszel no permite detectar ciertas formas de DIF** (especialmente DIF no uniforme).
 - **En el modelo de Rasch**, especialmente si se utiliza CML para estimar los parámetros de los ítems, los resultados para evaluar DIF en diferentes ítems son (más) **interdependientes**.

Por ejemplo, la presencia de ítems que perjudican el grupo focal conlleva, como artefacto, a que en otros ítems se presenta una ventaja para este grupo.

- → Recomendación: Utilizar ambos enfoques como métodos complementarios.

Mantel-Haenszel vs la TRI para detectar DIF

- Supongamos que, en el marco de la TRI, se ha ajustado el modelo de Rasch y se utiliza este modelo para detectar DIF.

En este caso, el principio para dividir la muestra total en “grupos de nivel” es el mismo: la puntuación total en el test.

- Al comparar ambos métodos, se destacan las siguientes diferencias:
 - A un nivel teórico-conceptual, con el método de Mantel-Haenszel **no se ha validado la puntuación total** como medida del nivel en el rasgo latente.
 - El método de **Mantel-Haenszel no permite detectar ciertas formas de DIF** (especialmente DIF no uniforme).
 - **En el modelo de Rasch**, especialmente si se utiliza CML para estimar los parámetros de los ítems, los resultados para evaluar DIF en diferentes ítems son (más) **interdependientes**.
Por ejemplo, la presencia de ítems que perjudican el grupo focal conlleva, como artefacto, a que en otros ítems se presenta una ventaja para este grupo.

- → Recomendación: Utilizar ambos enfoques como métodos complementarios.