

# Introducción a la medición en las ciencias sociales

Sesión 4 - Modelos de Análisis

Área de Investigación

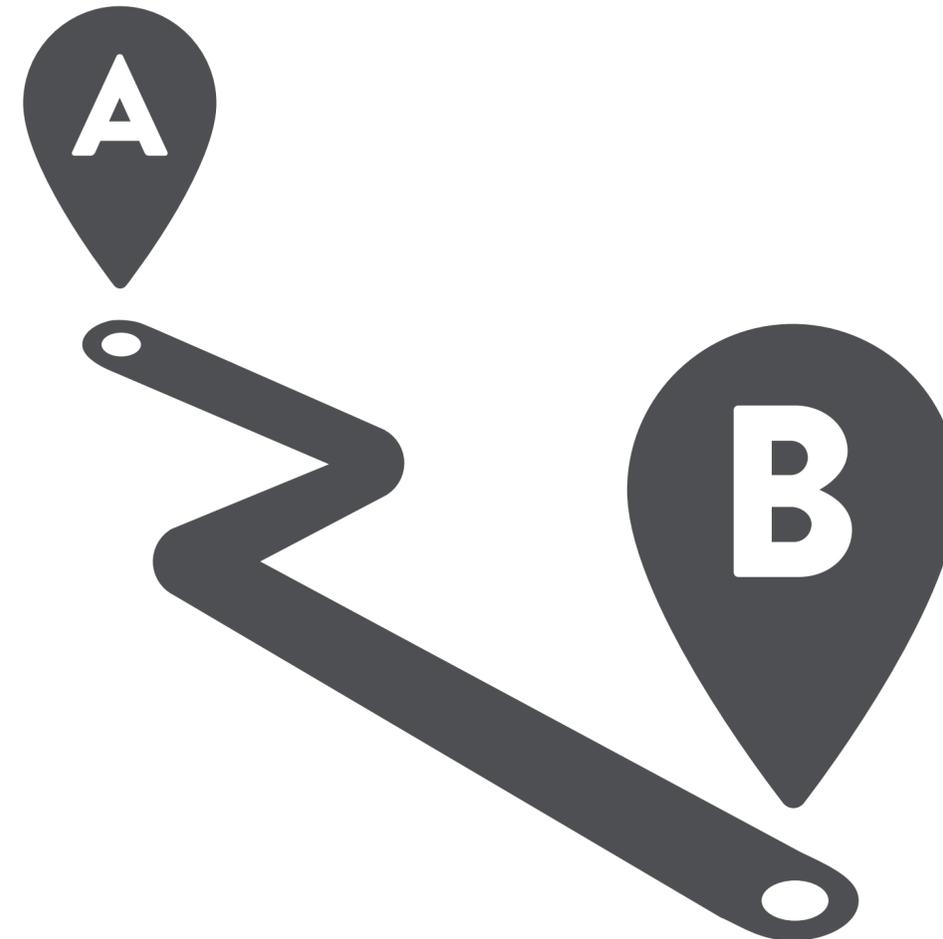


**Centro UC**  
Medición - MIDE

# Tabla de contenidos

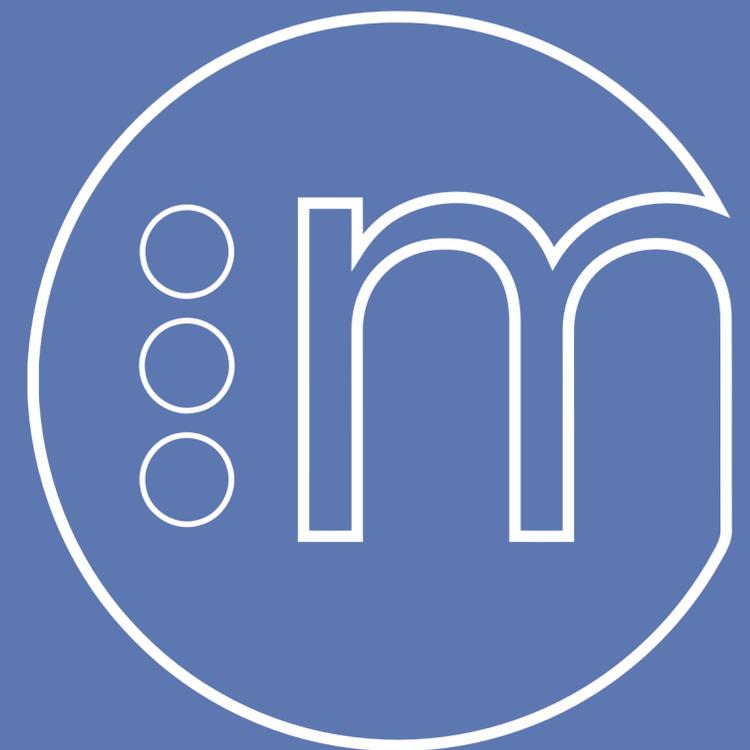
El plan para el día de hoy

- Modelamiento
- Modelos estadísticos para el análisis de respuestas
  - Teoría Clásica de Tests
  - Nexo entre puntajes verdaderos y variables latentes
  - Modelos de Variables Latentes
- Modelos de variables latentes para respuestas dicotómicas: El modelo de Rasch
- Marcos generales de modelos de variables latentes
- Problemas de Validez
- Resumen del proceso



# Dónde quedamos

... en las clases anteriores



# La lógica de las cuatro sesiones

Un mapa general de este curso

Sesión 1 — Medición: conceptos básicos

Sesión 2 — Validez, sesgo y confiabilidad

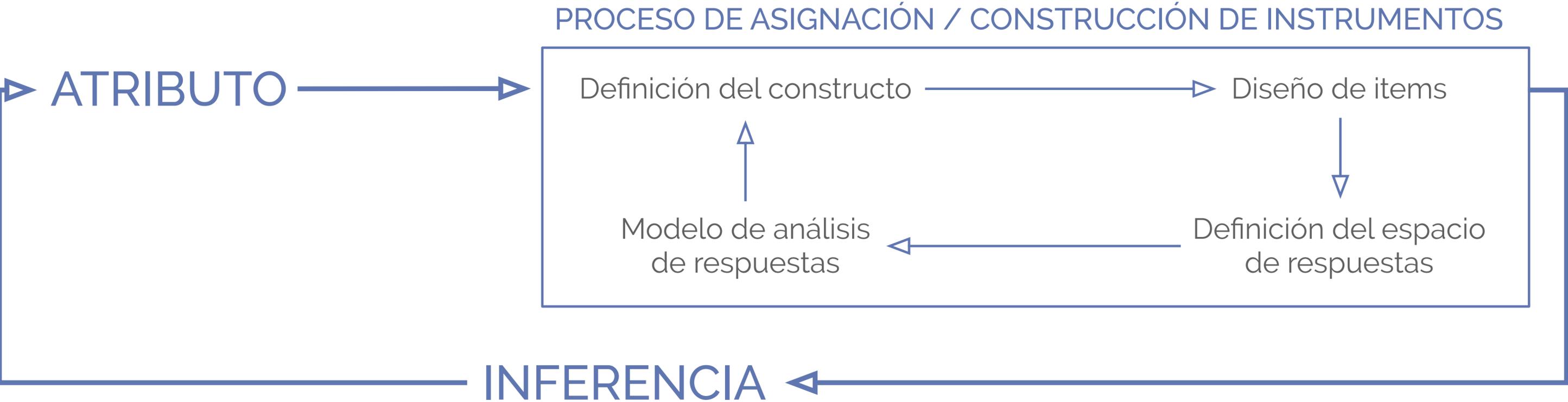
Sesión 3 — Construcción de instrumentos

Sesión 4 — Modelos de análisis

# Medición en ciencias sociales

Atributos, el proceso de asignación y las inferencias

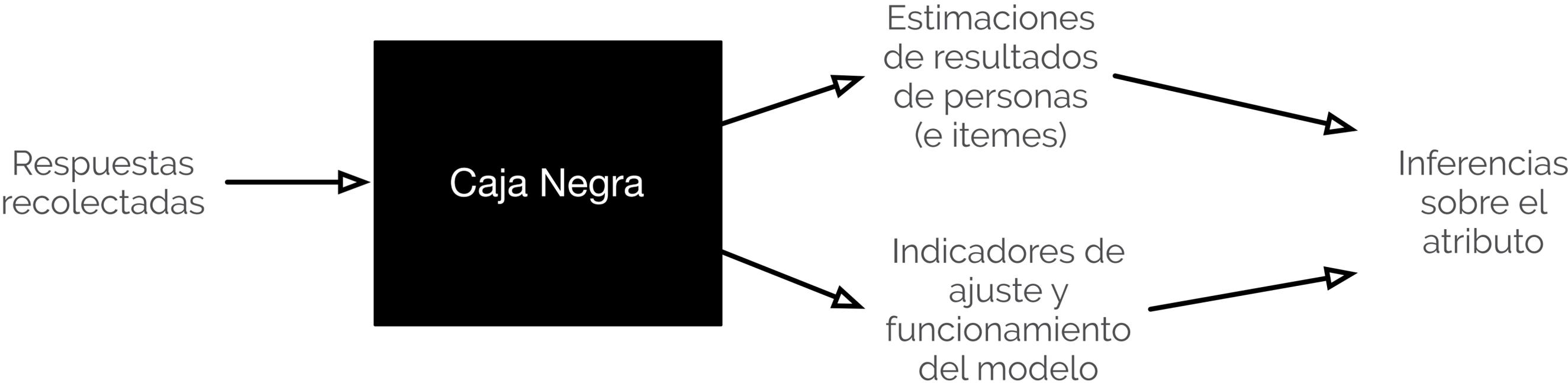
Recordemos nuevamente los tres elementos involucrados en el proceso de medición en las ciencias sociales:



# Modelamiento estadístico de las respuestas

La lógica general de la etapa de modelamiento estadístico

Hasta la clase de la semana pasada, los detalles respecto a los modelos se trataron como una caja negra.



# Modelamiento

... ¿qué son los modelos?



# ¿Qué es un modelo?

## Una definición

La idea de *modelo* es muy amplia, pero podemos entenderlos en general de la siguiente forma:



*Para un observador B, un objeto  $A^*$  es un modelo de un objeto A en la medida que B pueda usar  $A^*$  para responder preguntas que le interesen respecto a A*

— Minsky, 1965

Las ideas de los modelos y el modelamiento juegan un rol central en la investigación científica en general y en psicometría en particular.



Minsky, M. (1968). Matter, Mind and Models. In M. Minsky (Ed.), *Semantic information processing* (pp. 425–432).

¿Qué es un modelo?  
Veamos un ejemplo

# ¿Qué es un modelo?

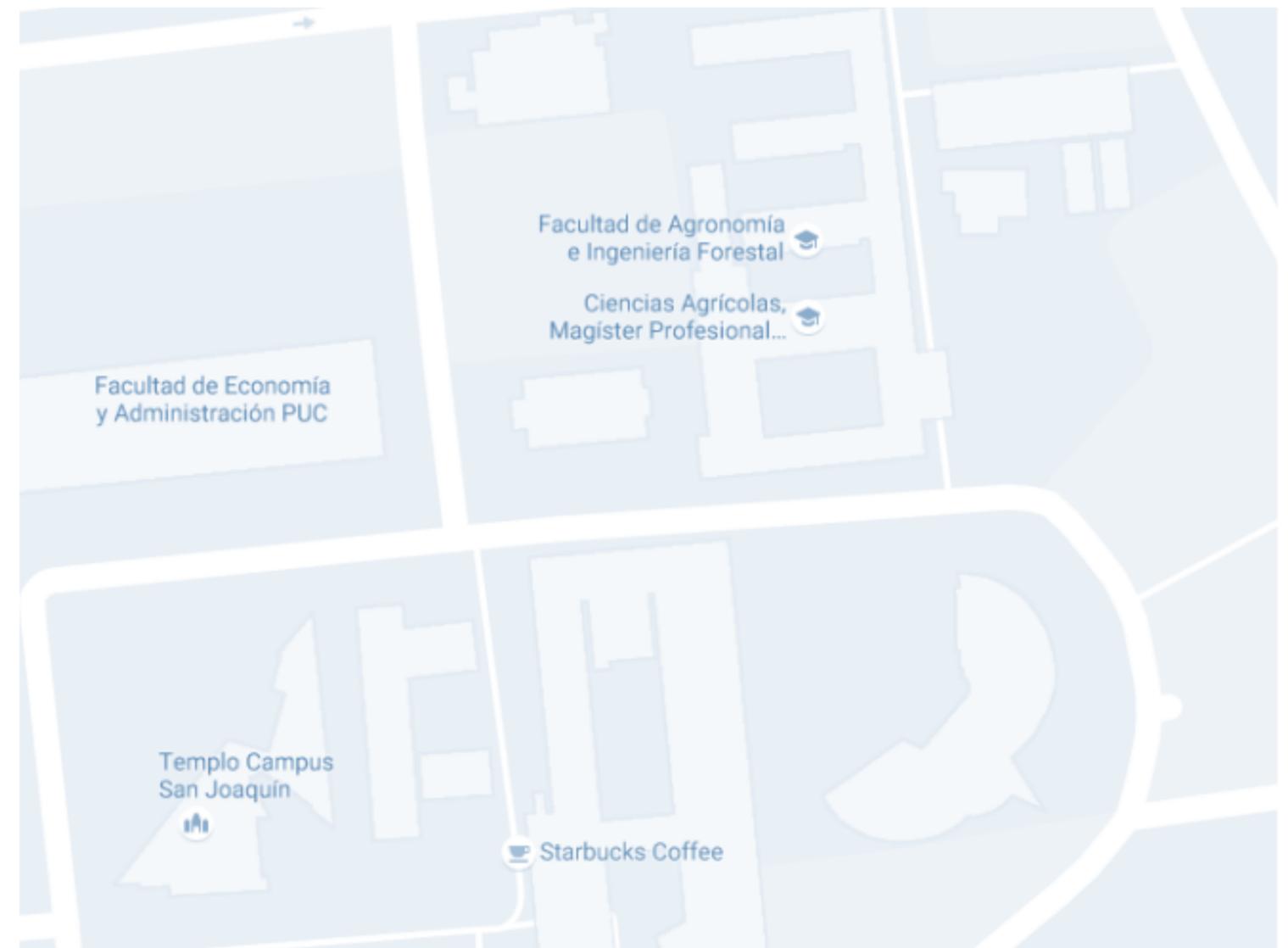
Una definición: ejemplo 1

Por ejemplo, un mapa es un modelo.

¿Alguien me puede decir, dónde está MIDE?

¿Por qué?

[pregunta retórica]



# ¿Qué es un modelo?

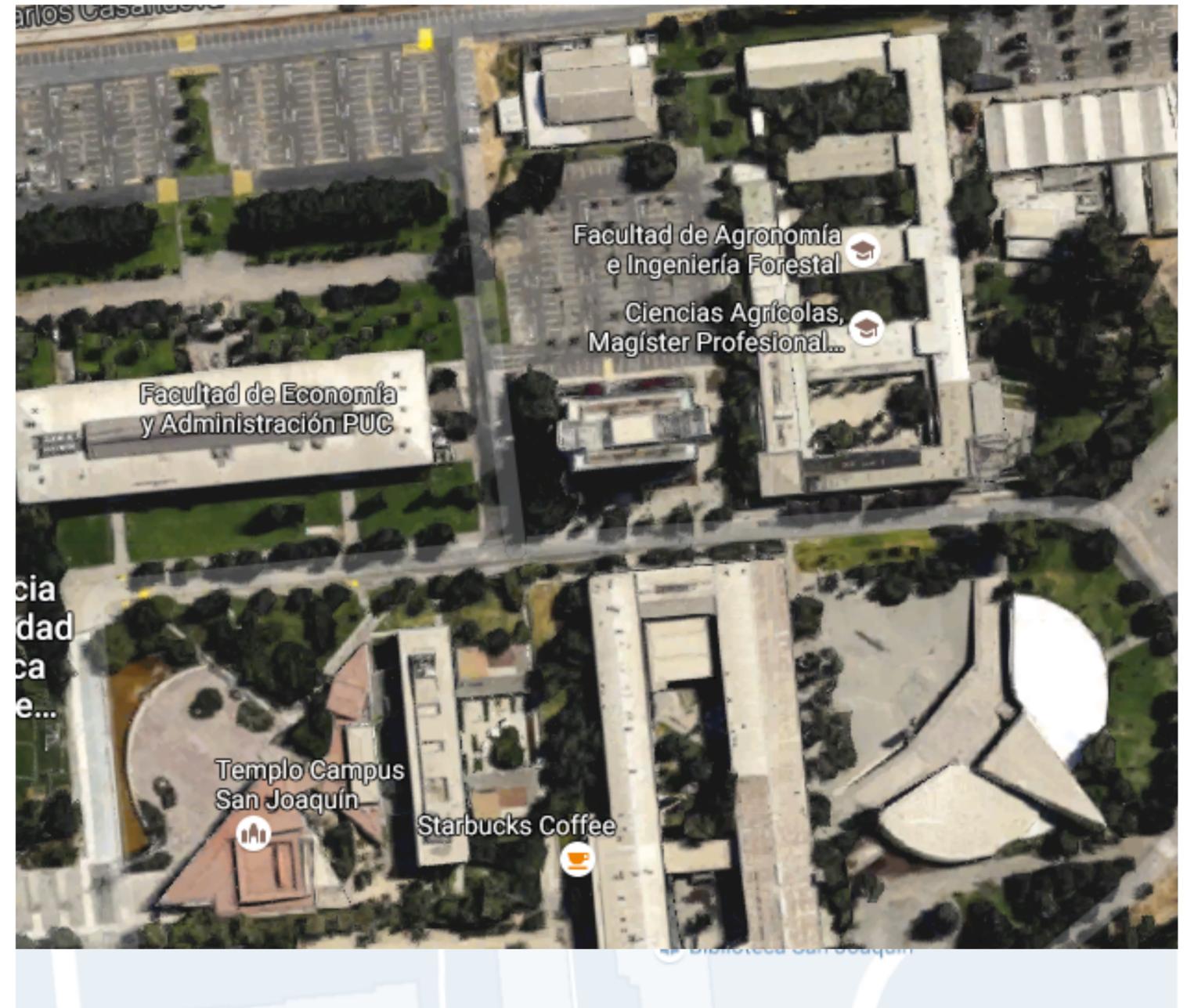
Una definición: ejemplo 1

Las mapas son representaciones a escala.

De objetos en un área.

De las distancias entre estos objetos.

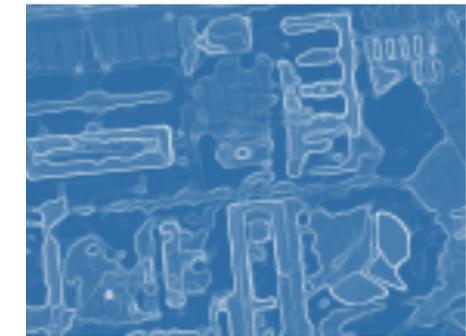
De la posición de estos en un espacio.



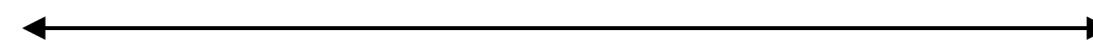
Un modelo, a pesar de ser una simplificación,  
puede ser suficientemente útil

# ¿Qué es un modelo?

Los modelos pueden diferir en fidelidad

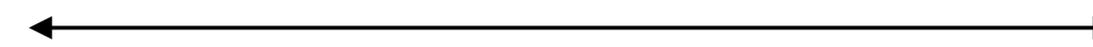


Mayor fidelidad



Menor fidelidad

Mayor ajuste



Menor ajuste

Así como en el caso de una imagen, podemos juzgar qué tan fiel representa lo fotografiado; en el caso de los modelos estadísticos, se puede evaluar qué tan fidedignos son a los datos. En general, a este aspecto, se le llama ajuste y se calculan diferentes indicadores de *ajuste de los modelos*.

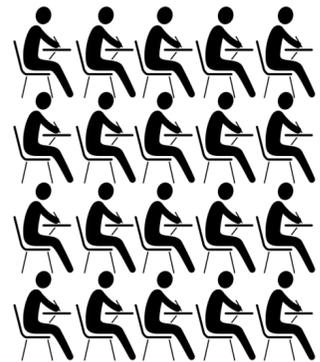


# ¿Qué estamos modelando?

## Relaciones entre variables

# ¿Qué es un modelo?

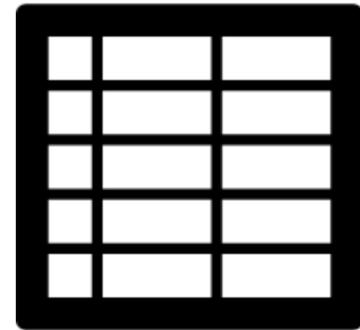
Lo modelado y sus representaciones



lo modelado

Un test de matemáticas

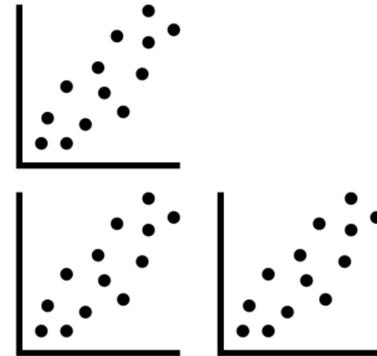
Escala de adhesión a la democracia



representación digital

base de datos en csv

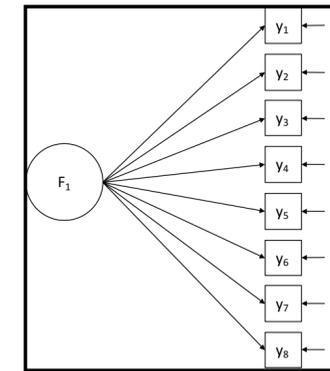
base de datos en excel



representación gráfica

*heatmap* de correlaciones

scatter matrix



representación diagramática

IRT

CFA

$$\begin{aligned}y_1 &= \lambda_1 F_1 + e_1 \\y_2 &= \lambda_2 F_1 + e_2 \\y_3 &= \lambda_3 F_1 + e_3 \\y_4 &= \lambda_4 F_1 + e_4 \\y_5 &= \lambda_5 F_1 + e_5 \\y_6 &= \lambda_6 F_1 + e_6 \\y_7 &= \lambda_7 F_1 + e_7 \\y_8 &= \lambda_8 F_1 + e_8\end{aligned}$$

representación formal

ecuaciones del modelo Rasch

ecuaciones de un CFA

Los diagramas que vamos a emplear del marco general de variables latentes, se encuentran en medio de una serie de otras representaciones sobre lo observado.

Existen varias formas de representar algo.

¿Qué modelos vamos a emplear?

Modelos de puntaje verdadero y modelos de variables latentes

# Modelos estadísticos para el análisis de respuestas



# Dos tradiciones de modelamiento

La teoría de puntaje verdadero y la teoría de respuesta al ítem

En psicología hay dos tradiciones principales sobre el análisis de respuestas a pruebas:

La **teoría clásica de test** y los **modelos de variables latentes**.

La primera se enfoca al modelamiento del puntaje verdadero de los tests.

La segunda se enfoca al modelamiento de variables latentes.

# Teoría Clásica de Tests

...o teoría clásica del puntaje verdadero



# La teoría clásica de tests

## Una introducción conceptual

La teoría clásica de tests—o teoría clásica del puntaje verdadero—es ampliamente utilizada y es consistente con la experiencia educacional respecto a como se asignan puntajes a una prueba.

La teoría clásica de tests (TCT) se focaliza en la idea de *puntaje verdadero*, en términos de que todo *puntaje observado* en una prueba es el producto de una combinación de dicho puntaje verdadero y un cierto grado de *error aleatorio*.

De acuerdo a esto, una ecuación básica de esta tradición es:

$$\text{Puntaje observado} = \text{Puntaje verdadero} + \text{Error}$$



Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, Mass: Addison-Wesley Pub. Co.  
Crocker, L. M., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart, and Winston.  
Hambleton, R. K., & Jones, R. W. (1993). An NCME instructional module on. *Educational measurement: issues and practice*, 12(3), 38-47.

# La teoría clásica de tests

## Una introducción conceptual

El puntaje verdadero de una persona puede ser definido como el promedio de todos los puntajes observados que se obtendrían **si pudiéramos repetir la misma prueba** a la misma persona y bajo las mismas condiciones.

Las inferencias en esta teoría son respecto al puntaje verdadero en un prueba en particular. Toda persona en tiene un determinado (y potencialmente distinto) puntaje verdadero en toda prueba.



Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, Mass: Addison-Wesley Pub. Co.  
Hambleton, R. K., & Jones, R. W. (1993). An NCME instructional module on. *Educational measurement: issues and practice*, 12(3), 38-47.

# La teoría clásica de tests

## Intuición principal

La ecuación básica de esta tradición se remonta a Edgeworth (1888), quien sugiere descomponer los puntajes observados, entre el puntaje verdadero, y un componente de error:

$$\text{Puntaje observado} = \text{Puntaje verdadero} + \text{Error}$$



Borsboom, D. (2005). *Measuring the Mind*. New York: Cambridge University Press.

# La teoría clásica de tests

## Ilustración del foco principal

Supongamos que conocemos el puntaje observado de un sujeto.

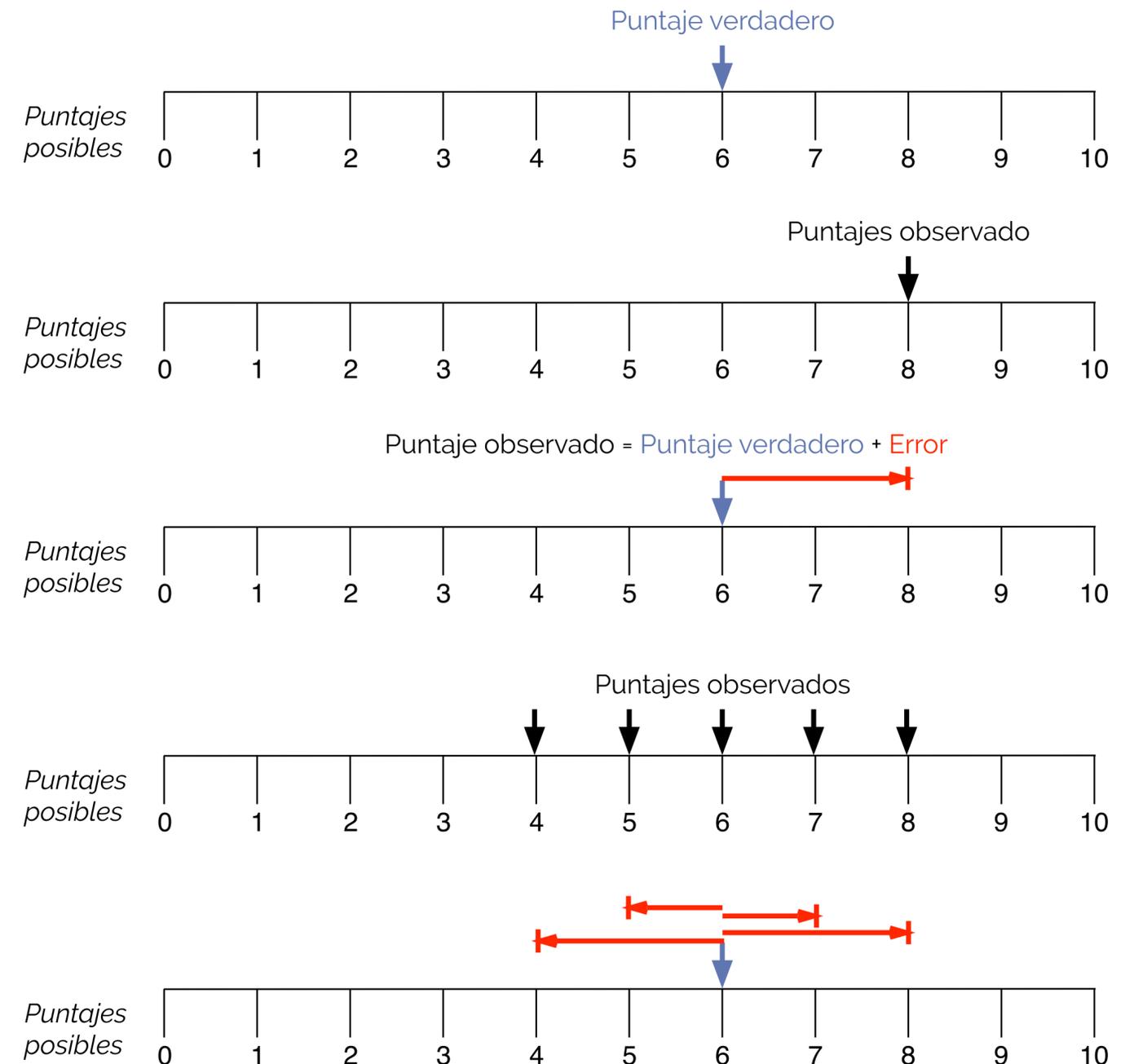
Sin embargo, en medidas subsecuentes, obtiene mayor puntaje, porque copió un par de respuestas.

**La distancia, entre su puntaje verdadero, y su observado constituye un error.**

Si realizamos cinco repeticiones adicionales, cuatro de éstas mostrarían diferentes distancias respecto del puntaje verdadero.

**Un promedio de todas estas mediciones sería nuestro estimador de cuál es el puntaje verdadero del sujeto.**

El puntaje verdadero puede ser definido como el promedio (valor esperado) de infinitas repeticiones de una prueba.



# Teoría clásica de tests

## Ideas centrales

El **puntaje verdadero** no es observable.

Usamos el **promedio** de todas las respuestas de una prueba, o su suma, como un **estimador** del puntaje verdadero.

Este estimador será mejor o peor dependiendo de los tamaños del error de la prueba. En otras palabras, qué tan “creíble” es mi estimador del puntaje verdadero depende de cuánto error de medición haya en la prueba aplicada.

Un supuesto respecto al error de medición en teoría clásica es que todas las mediciones que se realizan con una prueba particular comparten el mismo **error de medición estándar**.



Lord, F. M., & Novick, M. R. (1968). Statistical theories of mental test scores. Reading, Mass: Addison-Wesley Pub. Co.  
Crocker, L. M., & Algina, J. (1986). Introduction to classical and modern test theory. New York: Holt, Rinehart, and Winston.

# Teoría Clásica de Tests

Ejemplo: ICCS (2009) Ítems de conocimiento cívico, módulo latinoamericano, 16 ítems.



# Ejemplo: ICCS 2016

Marco de elementos a medir en el *International Civic and Citizenship Study 2016*

	Dominio de contenido 1: Sociedad cívica y sistemas	Dominio de contenido 2: Principios cívicos	Dominio de contenido 3: Participación cívica	Dominio de contenido 4: Identidad cívica
<b>Dominios cognitivos</b>				
<i>Conocimiento</i>	I	II	III	IV
<i>Razonamiento y aplicación</i>	V	VI	VII	VIII
<b>Dominios afectivos /conductuales</b>				
<i>Actitudes</i>	A	B	C	D
<i>Compromiso</i>	E	F	G	H

4 dominios de contenido: cada dominio de contenidos se desagrega en sub-dominios.

# Ejemplo: ICCS 2009

## Módulo Latinoamericano

Ítem ID	Formato	Clave	Punt. Máx.	Dominio	Sub Dominio	Contenido
LS2T06	Selección Múltiple	4	1	Principios cívicos	Libertad	Sin contenido especificado
LS2T13	Selección Múltiple	2	1	Sociedad cívica y sistemas	Ciudadanía	Derechos asignados y deseados por ciudadanos y grupos en la sociedad civil
LS2T01	Selección Múltiple	4	1	Principios cívicos	Equidad	Sin contenido especificado
LS2T14	Selección Múltiple	1	1	Sociedad cívica y sistemas	Ciudadano	Derechos asignados y deseados por ciudadanos y grupos en la sociedad civil
LS2T04	Selección Múltiple	4	1	Sociedad cívica y sistemas	Instituciones estatales	Sin contenido especificado

La "clave" del ítem LS2T06, fue corregida a lo que es expresado en la base de datos liberada.



Brese, F., Jung, M., Mirazchiyski, P., Schulz, W., & Zuehlke, O. (2014). ICCS 2009 User Guide for the International Database, Supplement 4 (2nd ed.). International Association for the Evaluation of Educational Achievement (IEA).

**Veamos algunos ejemplos  
de los ítems aplicados**  
**ICCS 2009, módulo latinoamericano**

# Ejemplo: Ítems ICCS 2009

Ítems liberados

LS2T02

**2** ¿Cuál de las siguientes es una manera democrática para que un gobierno trate a los grupos de la oposición?

- Permitirles expresar sus opiniones al público
- Obligarlos a cambiar de ideas y unirse al partido de gobierno
- Ofrecerles empleos en el gobierno
- Pedirles a líderes de otros países que hablen en contra de ellos



Brese, F., Jung, M., Mirazchyski, P., Schulz, W., & Zuehlke, O. (2014). ICCS 2009 User Guide for the International Database, Supplement 4 (2nd ed.). International Association for the Evaluation of Educational Achievement (IEA).

# Ejemplo: Ítems ICCS 2009

Ítems liberados

LS2T03

**3** ¿Cuál de las siguientes es una razón por la que existen las leyes de tránsito?

- Para fomentar el uso de transporte público
- Para recoger fondos destinados a pagar a la policía
- Para proteger la seguridad de los usuarios de calles y carreteras
- Para ayudar a planificar la construcción de calles y carreteras



Brese, F., Jung, M., Mirazchiyski, P., Schulz, W., & Zuehlke, O. (2014). ICCS 2009 User Guide for the International Database, Supplement 4 (2nd ed.). International Association for the Evaluation of Educational Achievement (IEA).

# Ejemplo: Ítems ICCS 2009

## Ítems liberados

LS2T06

- 6** ¿Cuál de las siguientes fue una de las consecuencias de las dictaduras latinoamericanas del siglo pasado?
- Se redujo significativamente la pobreza en los países en dictadura.
  - Numerosos nuevos inmigrantes se radicaron en los países en dictadura.
  - Muchos delincuentes comunes fueron liberados en los países en dictadura.
  - Muchos de los opositores tuvieron que salir de los países en dictadura.



Brese, F., Jung, M., Mirazchyski, P., Schulz, W., & Zuehlke, O. (2014). ICCS 2009 User Guide for the International Database, Supplement 4 (2nd ed.). International Association for the Evaluation of Educational Achievement (IEA).

# Ejemplo: Ítems ICCS 2009

Ítems liberados

LS2T07

7 ¿Cuál es el texto jurídico que establece los deberes y derechos individuales y regula el funcionamiento del Estado?

- El Código de Trabajo
- El Código de Derecho Civil
- La Constitución
- La Ley Internacional de Justicia



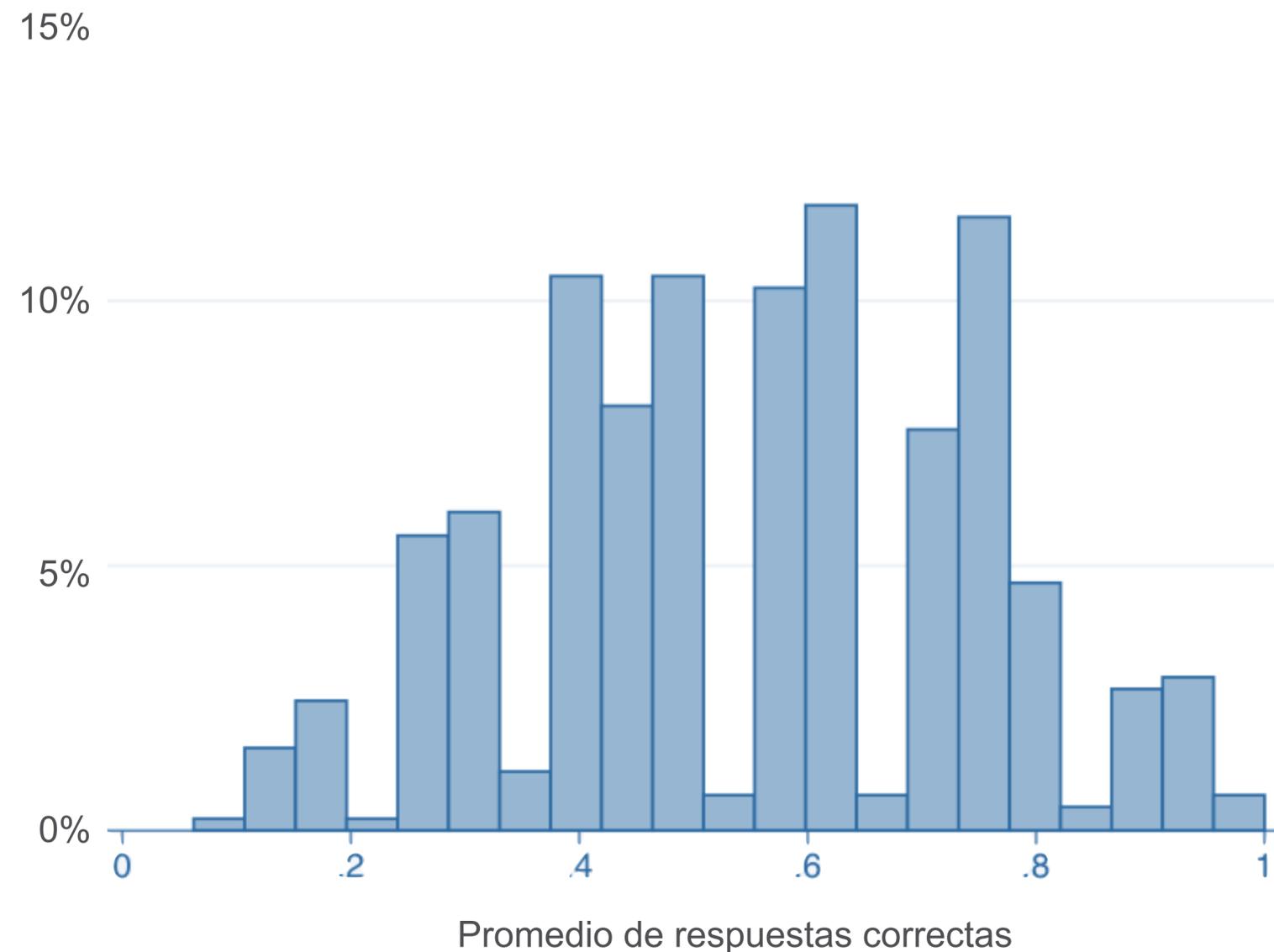
Brese, F., Jung, M., Mirazchyski, P., Schulz, W., & Zuehlke, O. (2014). ICCS 2009 User Guide for the International Database, Supplement 4 (2nd ed.). International Association for the Evaluation of Educational Achievement (IEA).

# Ejemplo: Ítems ICCS 2009

Podemos promediar las respuestas codificadas como correctas (valor 1) e incorrectas (valor 0) para obtener el puntaje observado de cada persona.

Este puntaje observado será nuestro **estimador del puntaje verdadero de cada persona**.

El gráfico de la derecha muestra la distribución de puntajes observados.



# Ejemplo: Ítems ICCS 2009

Puedo además, evaluar la relación entre las respuestas a cada pregunta con el puntaje total observado.

¿Qué pasa con el ítem seis?

code	d01	d02	d03	d04	d05	d06	d07	d08	d09	d10	d11	d12	d13	d14	d15	d16	mean_score
11060120	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	100.0%
10970203	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	93.8%
10880113	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	87.5%
11110114	1	1	1	1	1	0	1	1	0	0	1	1	1	1	1	1	81.3%
10260206	1	1	1	1	1	0	1	1	1	1	0	1	1	0	0	1	75.0%
11240211	1	1	1	1	1	0	1	1	1	0	0	0	1	1	1	1	75.0%
11760119	1	1	0	1	1	1	1	1	1	0	0	0	1	1	1	0	68.8%
12100105	1	1	1	1	1	0	1	0	1	0	1	1	0	0	1	1	68.8%
12160106	0	1	1	1	1	0	1	0	1	1	1	0	0	1	1	1	68.8%
11380106	0	1	1	1	0	0	1	1	1	1	0	1	1	0	0	1	62.5%
11550222	1	1	1	1	0	0	0	1	1	0	1	1	1	1	0	0	62.5%
11830207	0	1	0	1	1	0	1	0	0	0	1	1	1	0	1	1	56.3%
10070103	1	1	1	1	0	0	0	1	0	1	0	0	1	0	1	1	56.3%
10070125	0	1	1	1	0	0	1	1	1	1	1	0	0	0	1	0	56.3%
10240135	0	1	1	1	1	0	0	1	0	0	1	0	1	0	1	1	56.3%
10370511	0	1	1	1	0	0	1	1	0	0	1	1	0	1	0	0	50.0%
10840105	1	0	1	1	1	0	0	0	0	1	1	0	0	1	0	1	50.0%
11590427	0	1	0	1	1	0	0	0	1	0	0	0	1	1	1	1	50.0%
11670101	0	1	0	1	1	0	1	0	0	0	1	1	0	0	0	1	43.8%
10610506	0	0	1	1	1	0	0	0	1	0	1	0	1	0	1	0	43.8%
10970119	1	0	1	0	1	0	1	0	0	0	1	0	0	1	1	0	43.8%
10780234	1	1	0	0	0	1	1	0	0	0	0	0	1	0	0	1	37.5%
10320223	0	1	0	1	0	0	1	1	0	1	1	0	0	0	0	0	37.5%
11140619	1	1	1	1	0	0	0	1	0	1	0	0	0	0	0	0	37.5%
10060114	0	1	1	1	0	0	0	0	0	0	0	0	1	0	1	0	31.3%
11600106	1	0	1	0	1	0	0	0	0	0	0	0	1	1	0	0	31.3%
10630201	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0	1	25.0%
10710605	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0	1	25.0%
11190104	0	0	1	0	0	0	0	1	0	0	0	0	0	1	1	0	25.0%

# Ejemplo: Ítems ICCS 2009

“Dificultad” o valor  $p$ : proporción de correctas sobre el total de evaluados. En teoría clásica de tests a mayor índice de “dificultad” la pregunta es más fácil;

Proporción de incorrectas: proporción de evaluados que contesta incorrectamente cada ítem, es igual a  $1 - p$ ;

Correlación: correlación entre las respuestas del ítem y el puntaje total (en particular, esta es la correlación Biserial);

Alfa: Alfa esperado de la prueba, si el ítem es removido.

Ítems originales	Ítems corregidos	“Dificultad” valor $p$	Proporción de incorrectas	Correlación	Alfa
LS2T06 **	d06	0.24	0.76	0.25	0.68
LS2T12	d12	0.31	0.69	0.28	0.68
LS2T10	d10	0.38	0.62	0.38	0.67
LS2T07 **	d07	0.45	0.55	0.19	0.69
LS2T11	d11	0.45	0.55	0.41	0.67
LS2T13	d13	0.47	0.53	0.25	0.69
LS2T09	d09	0.48	0.52	0.35	0.67
LS2T01	d01	0.49	0.51	0.42	0.67
LS2T05	d05	0.57	0.43	0.25	0.69
LS2T14	d14	0.57	0.43	0.33	0.68
LS2T08	d08	0.63	0.37	0.44	0.67
LS2T16	d16	0.65	0.35	0.44	0.67
LS2T02	d02	0.69	0.31	0.41	0.67
LS2T15	d15	0.69	0.31	0.54	0.66
LS2T03	d03	0.81	0.19	0.56	0.66
LS2T04	d04	0.84	0.16	0.53	0.67

# Ejemplo: Ítems ICCS 2009

Ordenando las dificultades de los ítems podemos ver que el ítem 6, es el ítem más difícil de la prueba.

¿Es igual de difícil para todos los estudiantes?

Ítems originales	Ítems corregidos	"Dificultad" valor $p$	Proporción de incorrectas	Correlación	Alfa
LS2T06 **	d06	0.24	0.76	0.25	0.68
LS2T12	d12	0.31	0.69	0.28	0.68
LS2T10	d10	0.38	0.62	0.38	0.67
LS2T07 **	d07	0.45	0.55	0.19	0.69
LS2T11	d11	0.45	0.55	0.41	0.67
LS2T13	d13	0.47	0.53	0.25	0.69
LS2T09	d09	0.48	0.52	0.35	0.67
LS2T01	d01	0.49	0.51	0.42	0.67
LS2T05	d05	0.57	0.43	0.25	0.69
LS2T14	d14	0.57	0.43	0.33	0.68
LS2T08	d08	0.63	0.37	0.44	0.67
LS2T16	d16	0.65	0.35	0.44	0.67
LS2T02	d02	0.69	0.31	0.41	0.67
LS2T15	d15	0.69	0.31	0.54	0.66
LS2T03	d03	0.81	0.19	0.56	0.66
LS2T04	d04	0.84	0.16	0.53	0.67

# Ejemplo: Ítems ICCS 2009

## Ítems liberados

¿Cuál de las siguientes fue una de las consecuencias de las dictaduras latinoamericanas del siglo pasado?

Alternativa	Grupo bajo	Grupo medio	Grupo alto	Opciones de respuesta
1	0.248	0.268	0.216	Se redujo significativamente la pobreza en los países en dictadura
2	0.322	0.312	0.176	Numerosos nuevos inmigrantes se radicaron en los países en dictadura
3	0.315	0.217	0.184	Muchos de los delincuentes comunes, fueron liberados en los países en dictadura.
*4	0.114	0.203	0.424	Muchos de los opositores tuvieron que salir de los países en dictadura.

Los resultados, por tercios de “conocimiento cívico” (estimado en base al puntaje total) indican que es un ítem difícil, especialmente para los dos tercios inferiores. **¿Debiera eliminar el ítem? ¿o conservarlo?**

Notas: Los números de la tabla, corresponden a la tasa de respuesta frente a cada alternativa del ítem; estos están desplegados para 3 grupos de puntaje total, producto de dividir la distribución en tres partes iguales. **Alternativa** = refiere a las respuestas de los ítems de selección múltiple; **Grupo Bajo** = refiere al primer tercio de la distribución de puntajes totales; **Grupo Medio** = refiere al siguiente tercio de puntajes totales **Grupo Alto** = refiere al último tercio.

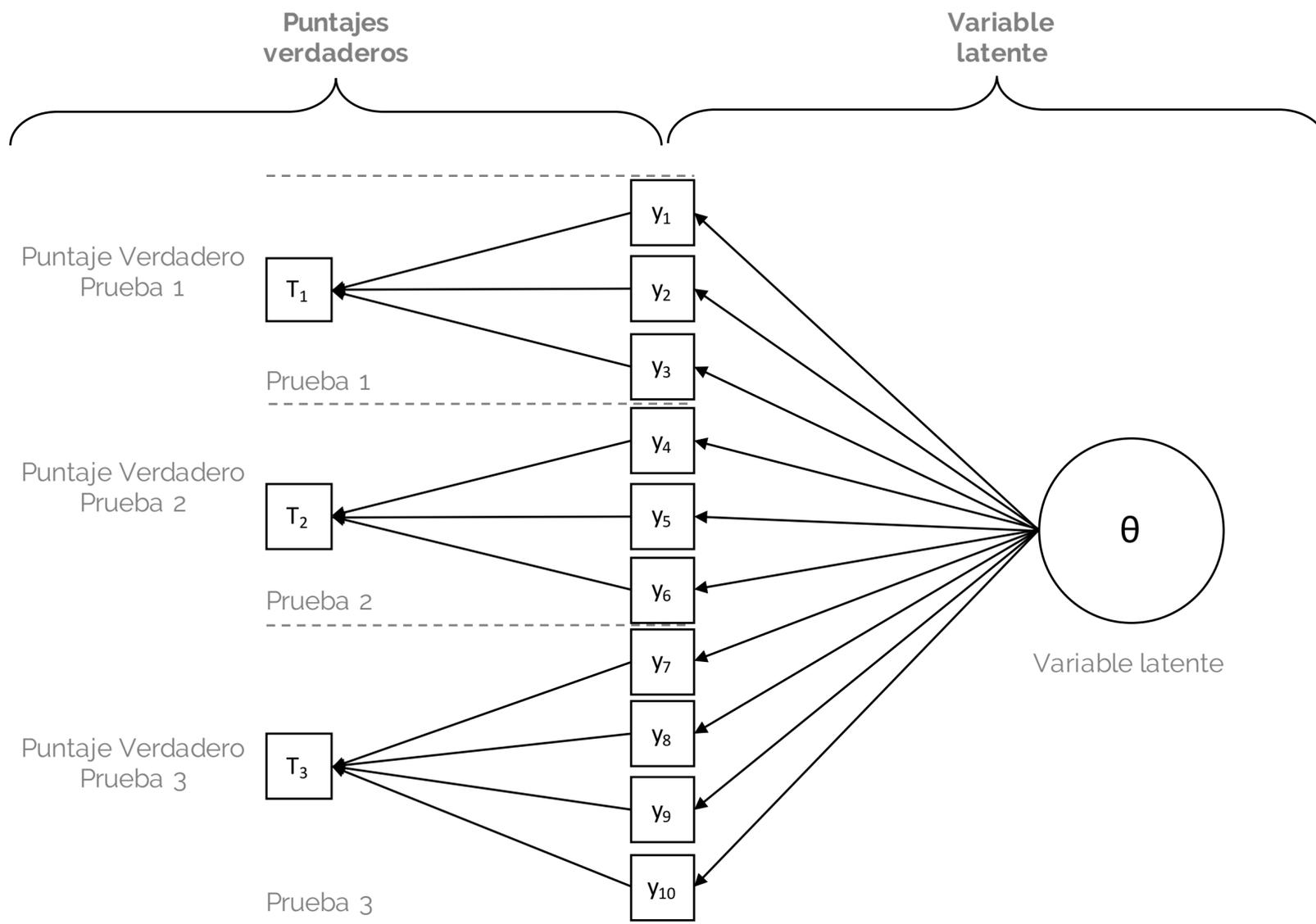
# Nexo entre puntajes verdaderos y variables latentes

¿cómo se vinculan estas dos tradiciones?



# Puntaje verdadero y variables latentes

## Nexo entre las tradiciones



El foco principal de teoría clásica de tests, es la estimación del puntaje verdadero.

Este puntaje es **específico a cada test** y es diferente de la noción de un atributo latente.

Estos dos conceptos, si bien están relacionados, son distintos.

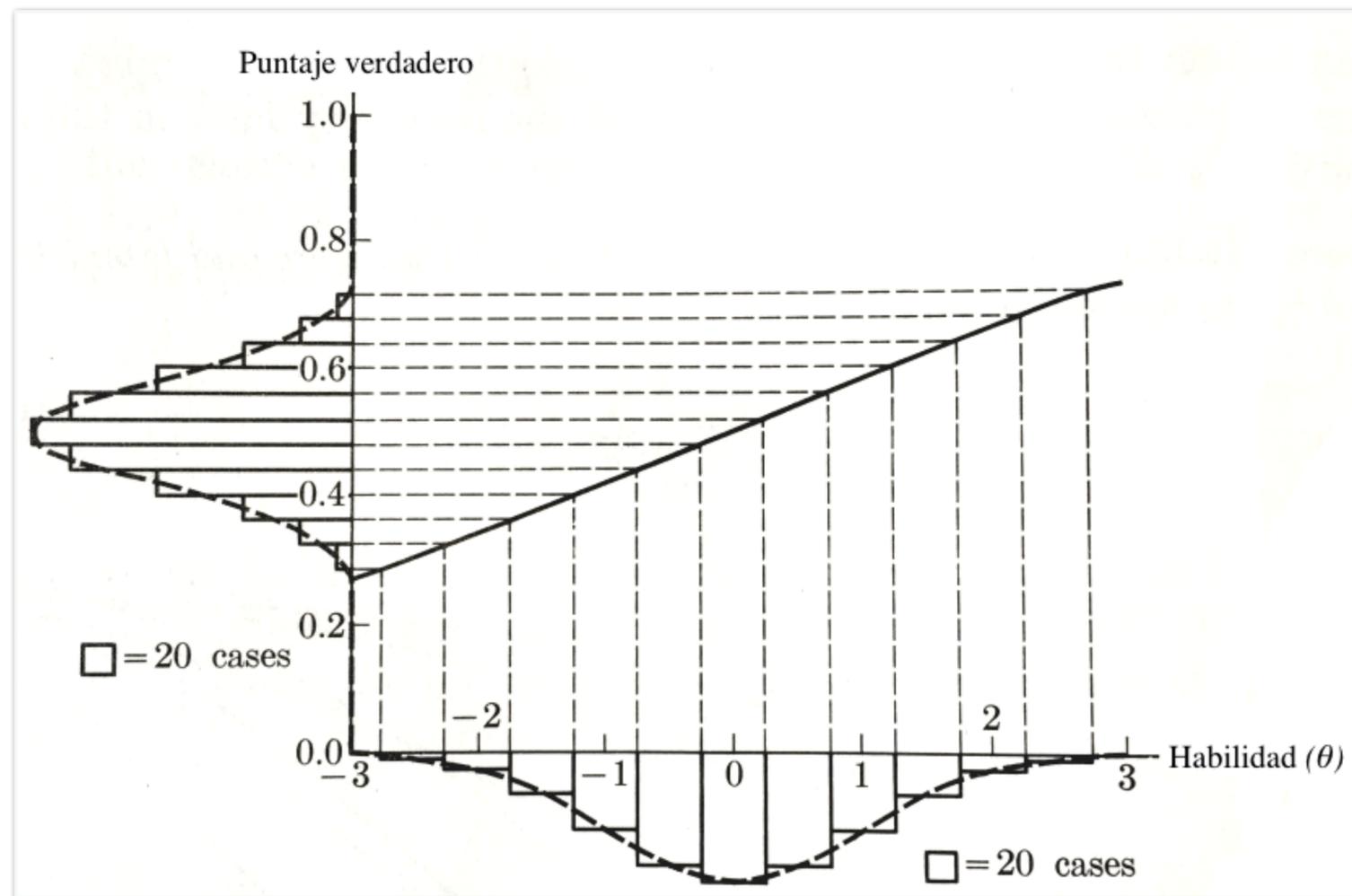
El primero concierne particularmente al puntaje esperado que se obtendría al responder un conjunto particular de preguntas.

La idea de variable latente, es más general, y se plantea para explicar las variaciones de respuesta, no solo de una prueba específica, sino de toda prueba que mida un mismo atributo.



# Puntaje verdadero y variables latentes

Nexo entre las tradiciones



Lord & Novick (1968; p 388)

Lord & Novick (1968) ilustran la relación entre estos dos conceptos. Es la “habilidad” o factor latente, la que explica los puntajes verdaderos esperados.

La idea de postular un factor latente que da cuenta de la relación entre variables observadas se remonta al trabajo de Spearman (1904).

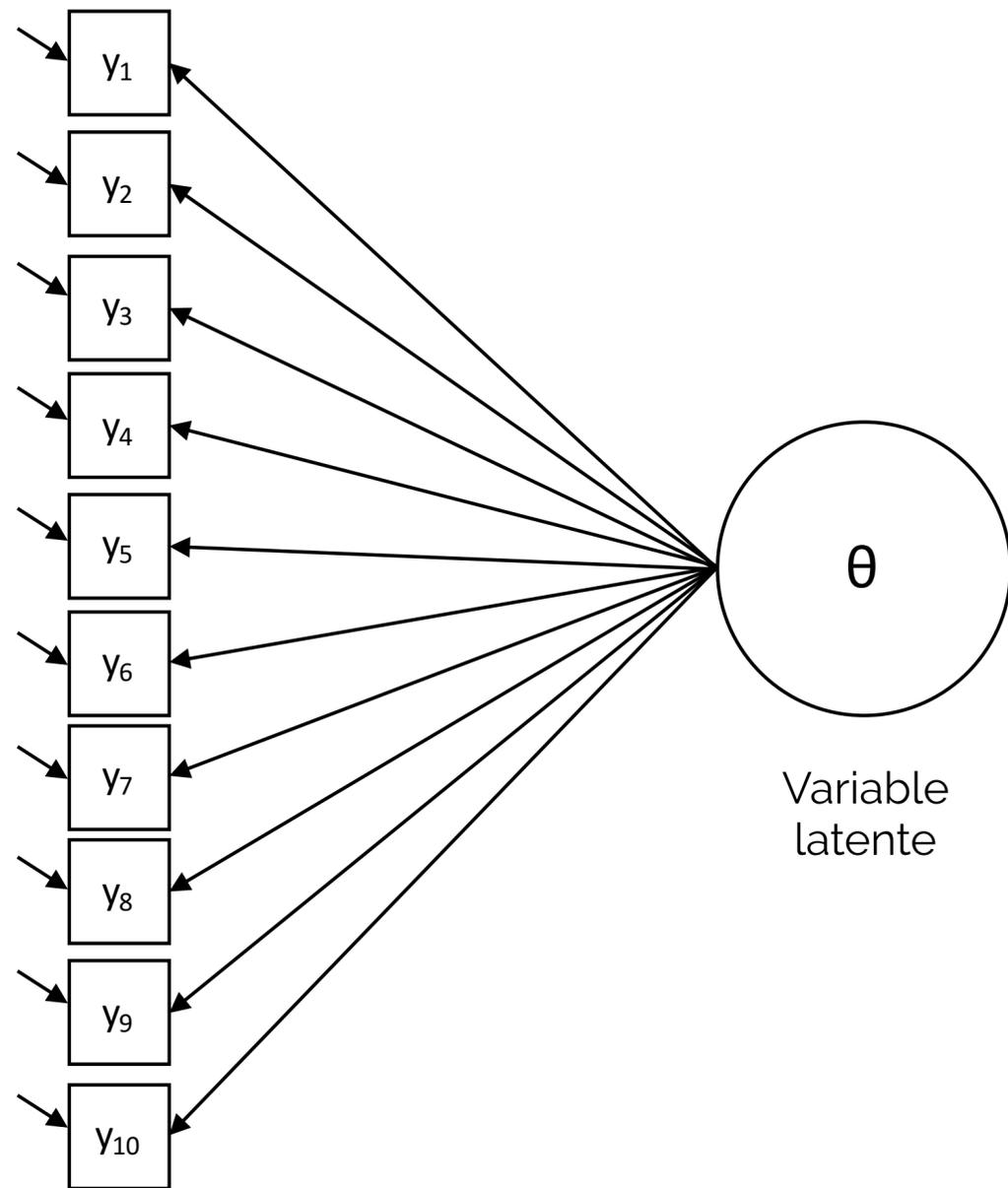
La idea general es que un conjunto de indicadores de una atributo (i.e. respuestas a preguntas en una prueba) puede ser descompuesto en un **factor general** y un **factor específico**. Estos dos factores explicarían los puntajes obtenidos en un test.



Lord, F. M., & Novick, M. R. (1968). Statistical theories of mental test scores. Reading, Mass: Addison-Wesley Pub. Co.  
Spearman (1904) "General intelligence," objectively determined and measured, American Journal of Psychology 15, 201-293.

# Puntaje verdadero y variables latentes

## Nexo entre las tradiciones



La idea de un factor general sería formalizada por **Thurstone (1947)**, bajo la idea del modelo de factor común (“**common factor model**”).

En este, se plantea que los puntajes observados de un indicador son explicados por una combinación lineal entre uno o más factores comunes y un factor único.

Al igual que en el método anterior, la idea es descomponer los puntajes observados en componentes. En este caso, la varianza común y la varianza única.

La primera, consiste en la variación explicada por el factor latente. La segunda es la varianza particular del indicador y error aleatorio (i.e. error de medición). Este es uno de los modelos en el enfoque de variables latentes.



Thurstone, L. L. (1947). Multiple factor analysis. Chicago, IL: The University of Chicago Press.

Brown, T. A. (2006). Confirmatory factor analysis for applied research. New York and London: Guilford Press.

Thorndike, R. M. (2005). History of Factor Analysis: A Psychological Perspective. In Encyclopedia of Statistics in Behavioral Science (Vol. 2, pp. 842–851). Chichester, UK: John Wiley & Sons, Ltd.

# Puntaje verdadero y variables latentes

## Modelos estadísticos

La invitación en este momento de la clase, es a pensar el modelamiento de respuestas, desde un marco general de variables latentes.

Para poder realizar esto, necesitamos introducir, algunas convenciones sobre el uso de diagramas que comúnmente se emplean en este marco.

Una vez que compartamos este marco de referencia, podemos ilustrar las formas que poseen determinados modelos estadísticos.

# Modelos de Variables Latentes



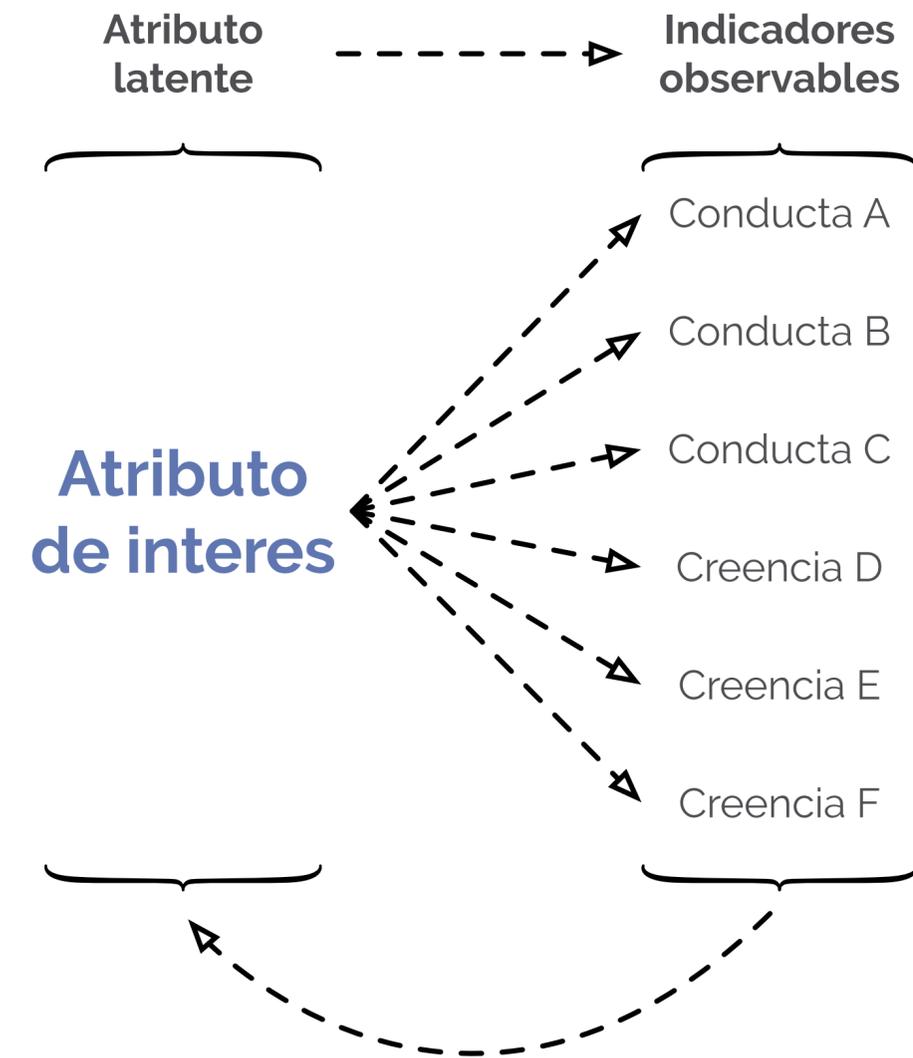
# Atributos latentes

## Los atributos en las ciencias sociales

Como se mencionó en clases anteriores, creemos que muchos de los atributos que queremos medir en ciencias sociales **no son observables directamente pero se manifiestan a través de indicadores visibles.**

Las respuestas a las condiciones estandarizadas que presentan los ítems serán los indicadores que usaremos para hacer inferencias respecto al atributo.

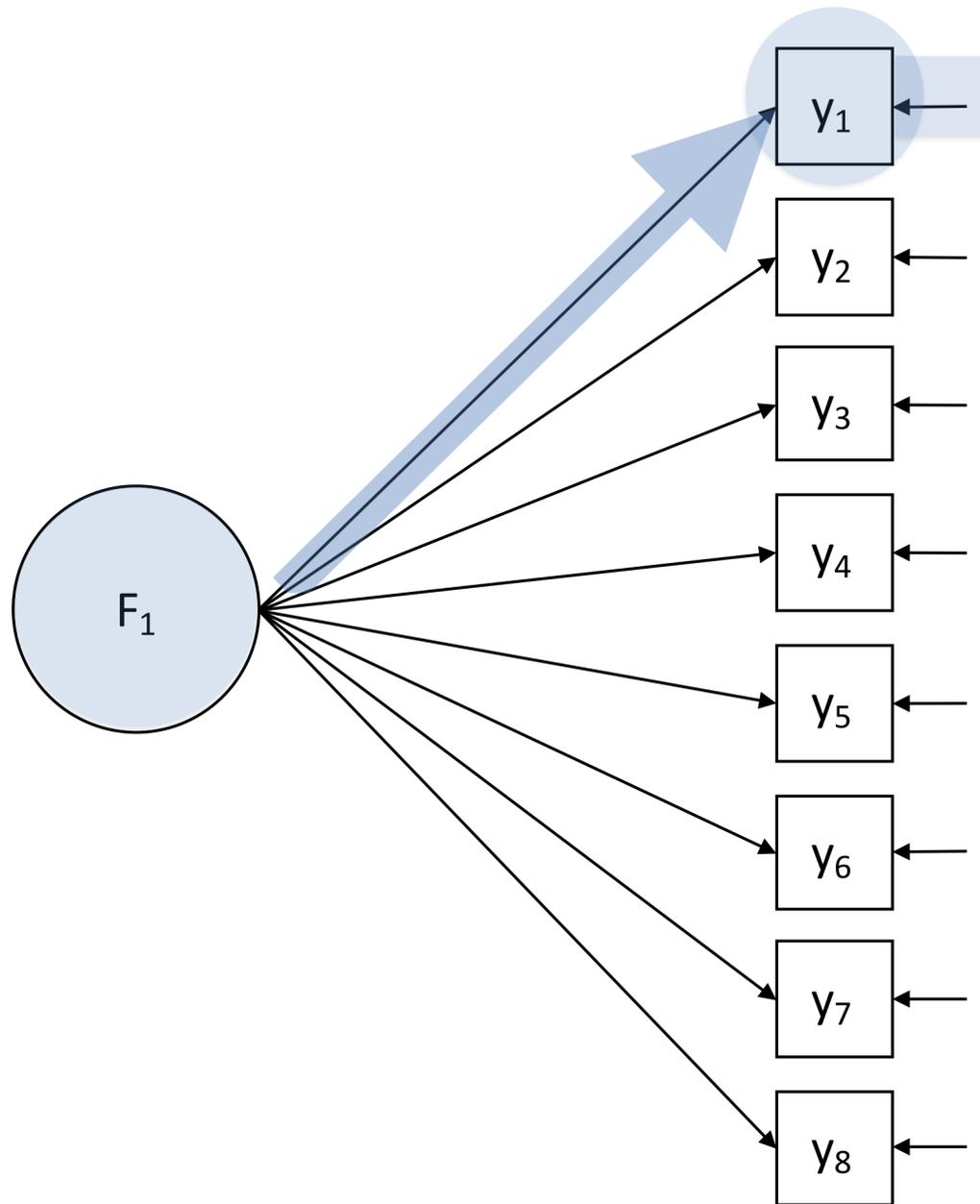
Hipotetizamos que el atributo latente explica cambios en indicadores observables.



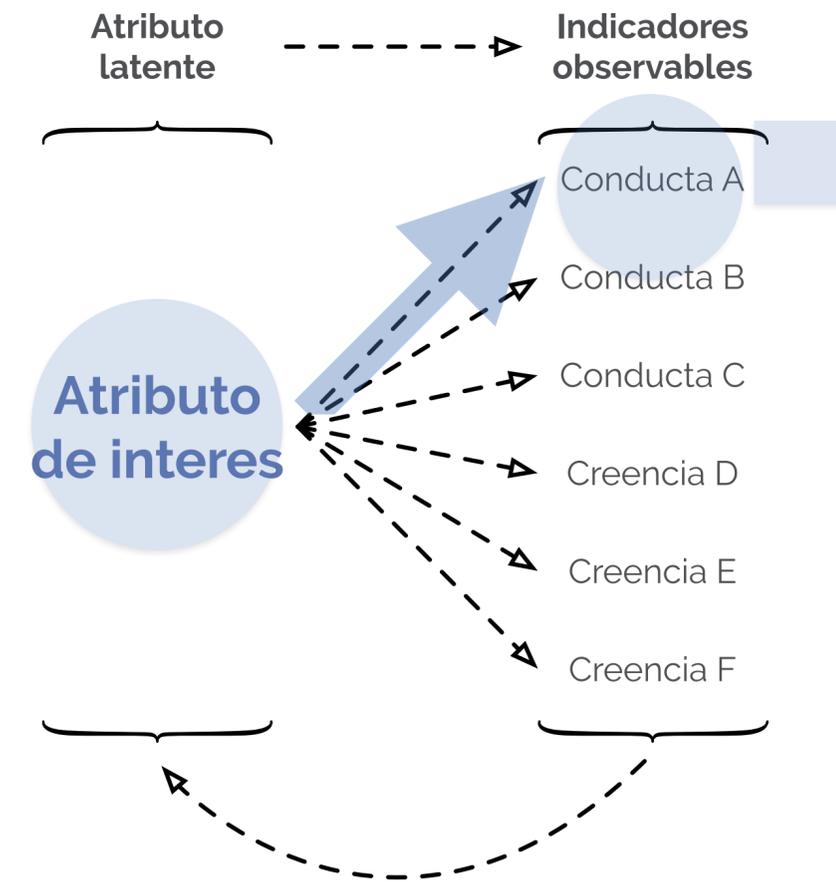
Dado esa hipótesis, hacemos inferencias sobre el atributo latente en base a esos indicadores.

# Modelos de variables latentes

Diagramas (o grafos) acíclico dirigidos para expresar modelos de variables latentes



Hipotetizamos que el atributo latente explica cambios en indicadores observables.



Dado esa hipótesis, hacemos inferencias sobre el atributo latente en base a esos indicadores.



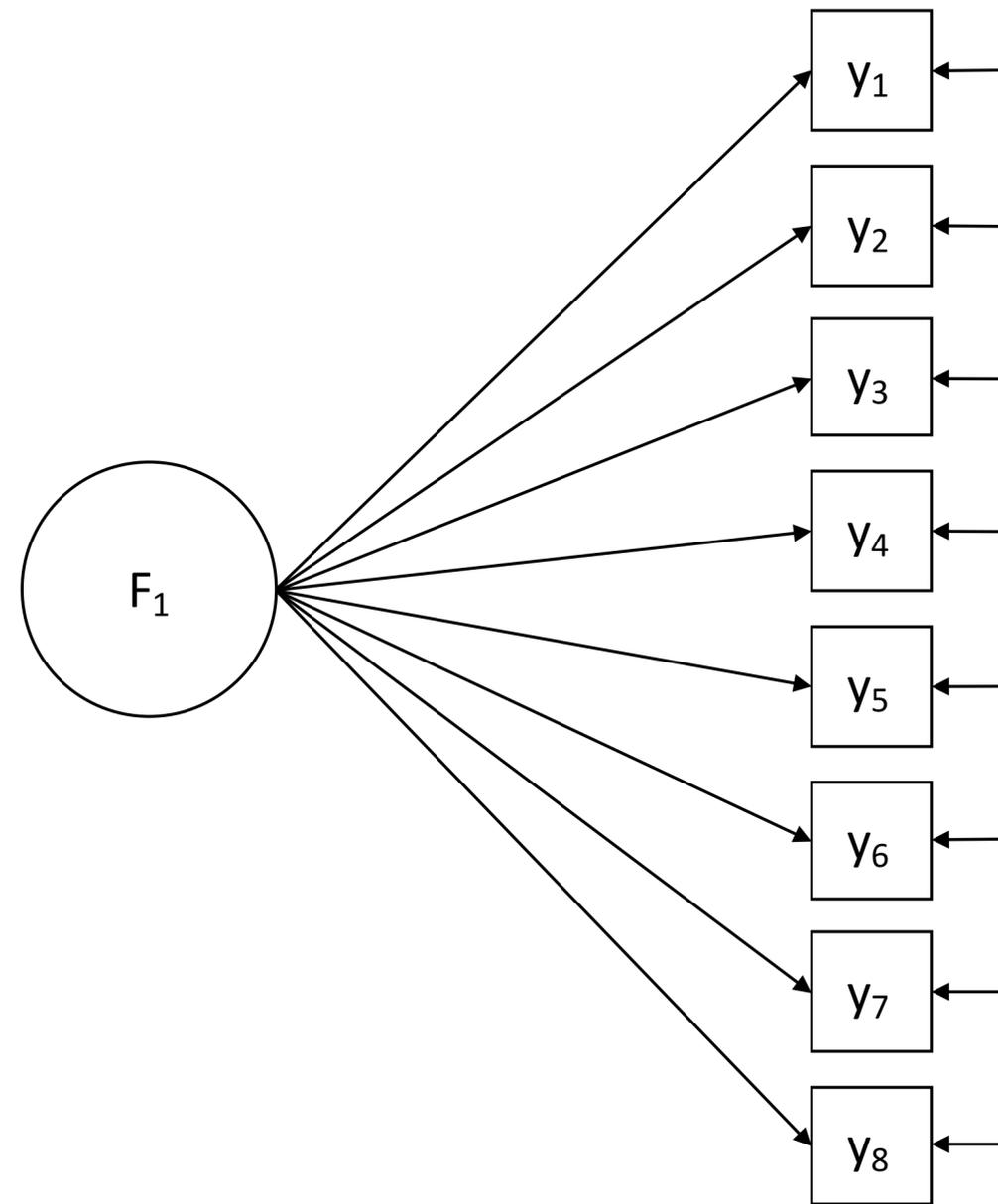
# Modelos de variables latentes

Diagramas (o grafos) acíclico dirigidos para expresar modelos de variables latentes

Este mismo modelo puede ser representado mediante **ecuaciones**, las cuales serían la **representación matemática** de estos modelos estadísticos.

No obstante, emplearemos los **diagramas** de los modelos para ilustrar su forma.

Usamos estos diagramas para expresar cómo están especificados estos modelos estadísticos.



$$y_1 = \lambda_1 F_1 + e_1$$

$$y_2 = \lambda_2 F_1 + e_2$$

$$y_3 = \lambda_3 F_1 + e_3$$

$$y_4 = \lambda_4 F_1 + e_4$$

$$y_5 = \lambda_5 F_1 + e_5$$

$$y_6 = \lambda_6 F_1 + e_6$$

$$y_7 = \lambda_7 F_1 + e_7$$

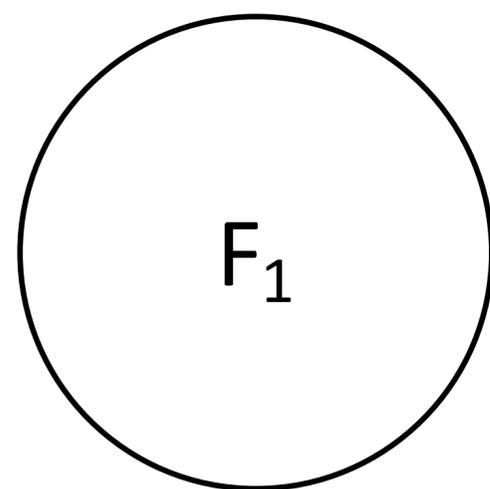
$$y_8 = \lambda_8 F_1 + e_8$$



Veamos los elementos  
que componen los diagramas.  
Revisémoslos uno a uno.

# Diagramas de variables latentes

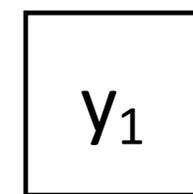
## Componentes



Variable latente  
no observada



Relación/Efecto



Variable  
observada

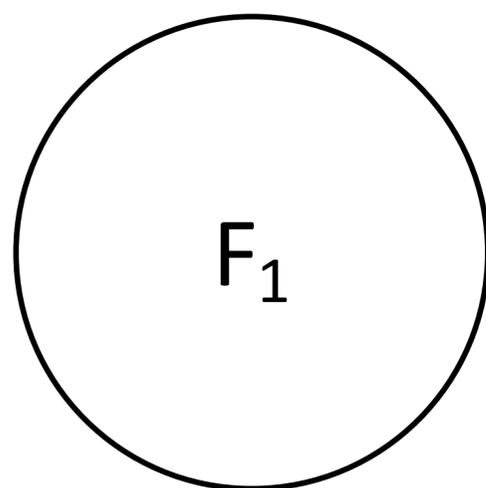


Residual



# Diagramas de variables latentes

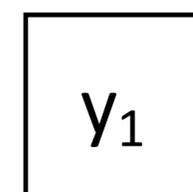
## Componentes



Variable latente  
no observada

Esta no existe en la base de datos.

Esta variable se “compone” de las relaciones entre variables observadas.



Variable  
observada

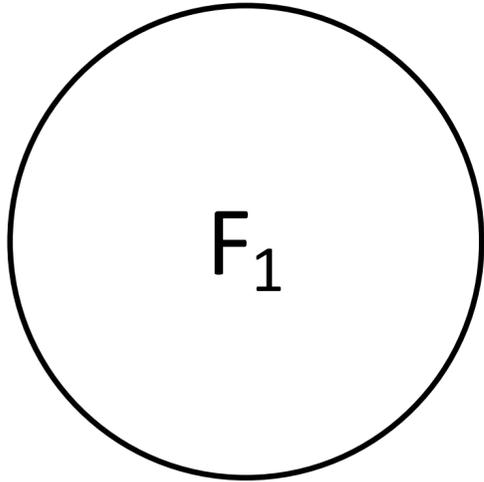
Esta variable se encuentra en la base de datos.

Comúnmente le llamamos indicador o ítem.



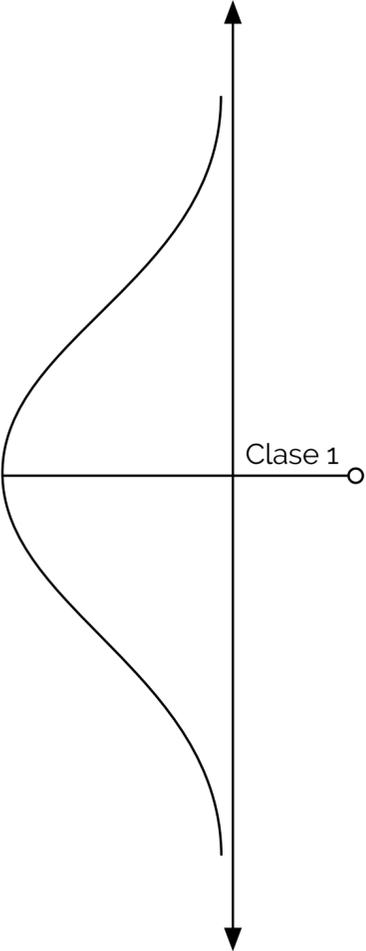
# Diagramas de variables latentes

Componentes: Variables latentes

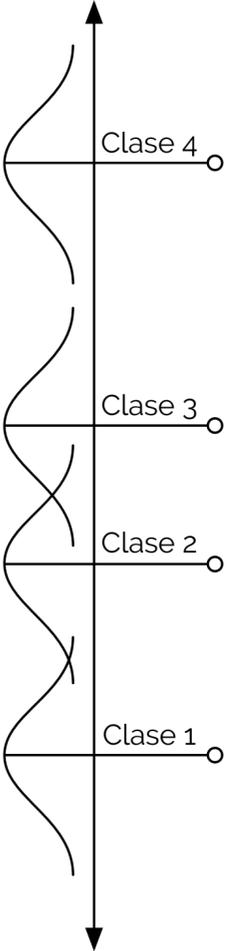


Variable latente

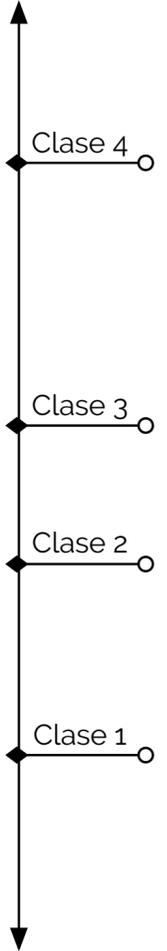
Estructura Cuantitativa +-----+ Estructura Ordinal +-----+ Estructura Categorical



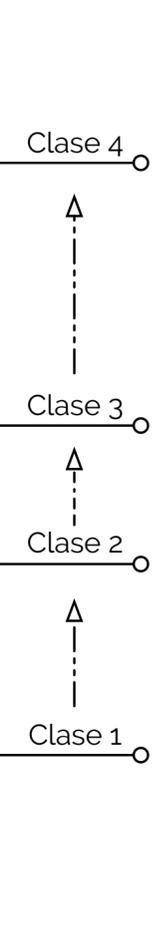
(a) Variable Cuantitativa



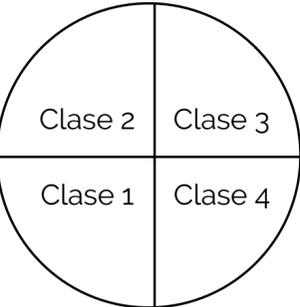
(b) Clases Heterogéneas Localizadas



(c) Clases Homogéneas Localizadas



(d) Clases Ordenandas

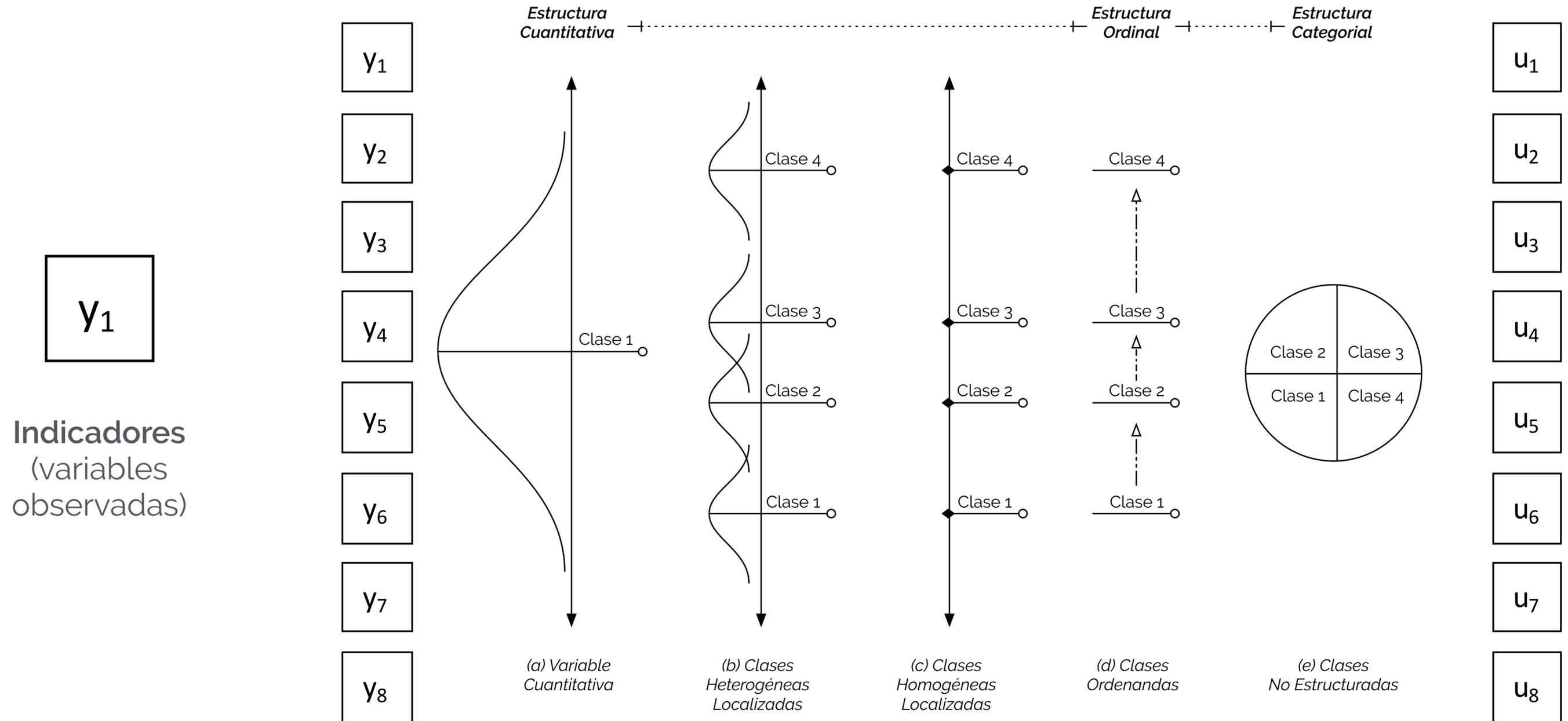


(e) Clases No Estructuradas

Dentro de este marco general, las variables latentes pueden adoptar diferentes tipos de estructura; desde estructuras continuas, a estructuras ordinales y categóricas.

# Diagramas de variables latentes

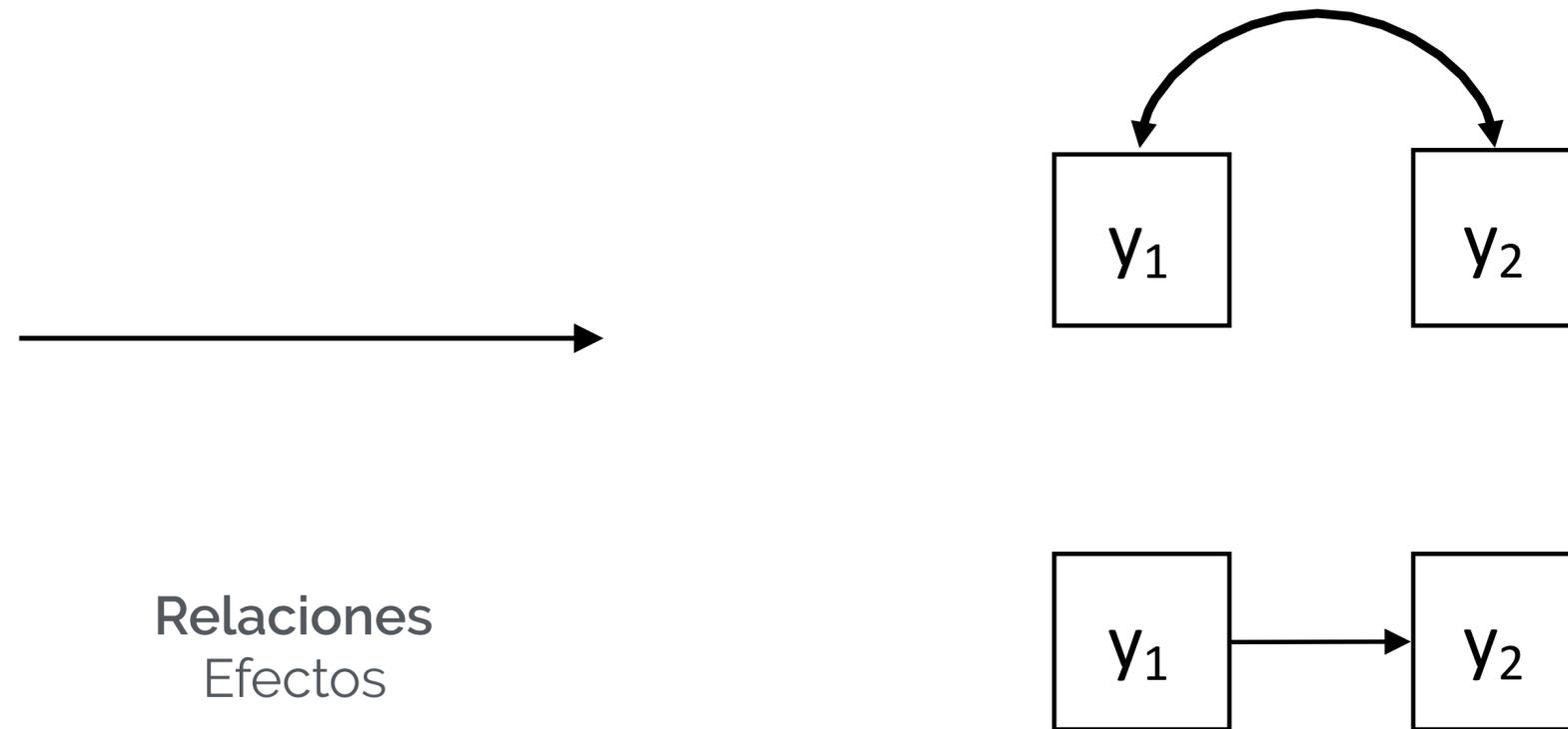
Componentes: Indicadores



Dentro del marco general de variables latentes, los indicadores pueden ser continuos ( $y_1$ - $y_6$ ); dicotómicos ( $u_1$ - $u_6$ ); y otra tipo de indicadores (e.g. con distribución de "zero-inflated-poisson").

# Diagramas de variables latentes

## Componentes: Relaciones



Relaciones  
Efectos

Dos grandes tipos de relaciones: líneas curvas y líneas rectas.

Líneas curvas expresan relaciones bidireccionales entre variables, por ejemplo, correlaciones. En general se interpreta como “ $y_1$  esta relacionado a  $y_2$ ”

Líneas rectas, expresan relaciones dirigidas entre variables, las que incluso podrían ser interpretadas causalmente. En general, se interpreta como “ $y_1$  explica a  $y_2$ ”.

Aparte de expresar las relaciones entre las variables latentes y los indicadores, también se pueden expresar relaciones entre indicadores. Además, estas relaciones pueden ser de diferente forma. Estas mismas relaciones se podrían establecer entre variables latentes.



Pearl, J. (1998). Graphs, causality, and structural equation models. *Sociological Methods & Research*, 27(2), 226-284.

Stata. (2013) Learning the language: Path diagrams and command language. In *Stata Structural Equation Modeling Reference Manual*. College Station, Texas: Stata Press

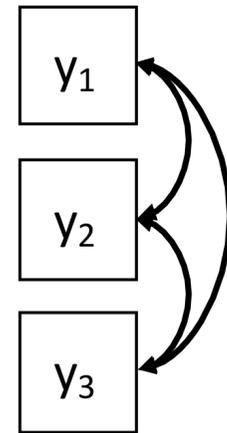
Tu, Y. K. (2012). Directed Acyclic Graphs and Structural Equation Modelling. In *Modern Methods for Epidemiology* (pp. 191-203). Springer Netherlands.

# Diagramas de variables latentes

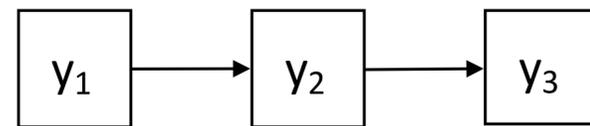
## Componentes: Relaciones



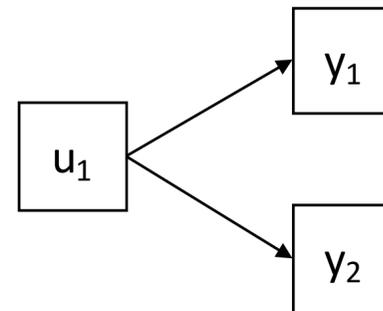
Relaciones  
Efectos



Líneas curvas expresan correlaciones entre variables. Estos son un conjunto de ítems correlacionados entre sí.



Las variables podrían estar secuencialmente relacionadas.



Una misma variable podría estar relacionada a varios indicadores.



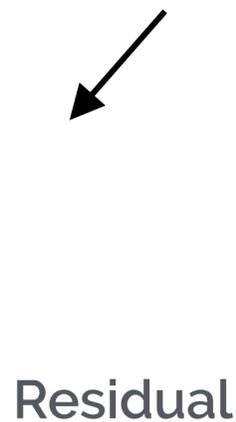
Pearl, J. (1998). Graphs, causality, and structural equation models. *Sociological Methods & Research*, 27(2), 226-284.

Stata. (2013) Learning the language: Path diagrams and command language. In *Stata Structural Equation Modeling Reference Manual*. College Station, Texas: Stata Press

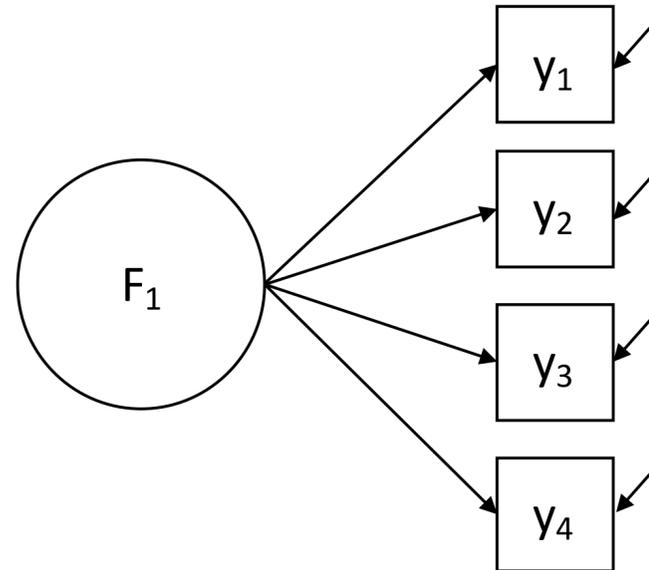
Tu, Y. K. (2012). Directed Acyclic Graphs and Structural Equation Modelling. In *Modern Methods for Epidemiology* (pp. 191-203). Springer Netherlands.

# Diagramas de variables latentes

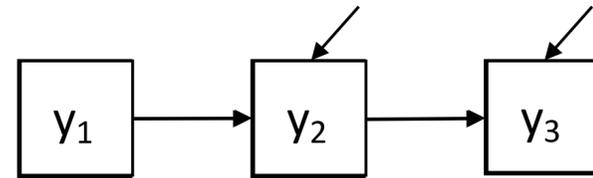
## Componentes: Residual



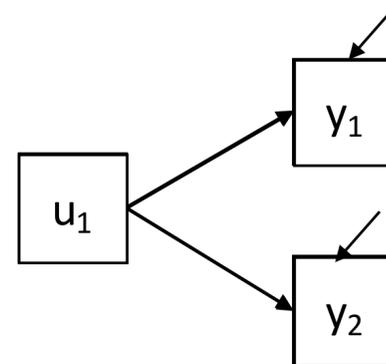
Los residuales, se agregan cada vez que la relación entre una variable y otra implica que se estime un error de predicción. Indicarlos es importante porque también existen modelos donde se fijan en cero.



Cuando una variable latente predice diferentes indicadores, podemos estimar o especificar residuales. Estos son empleados para contabilizar el error de medición.



Lo mismo sucede, cuando establecemos relaciones secuenciales entre variables.



O cuando tenemos una sola variable observada que predice a otras.



Pearl, J. (1998). Graphs, causality, and structural equation models. *Sociological Methods & Research*, 27(2), 226-284.

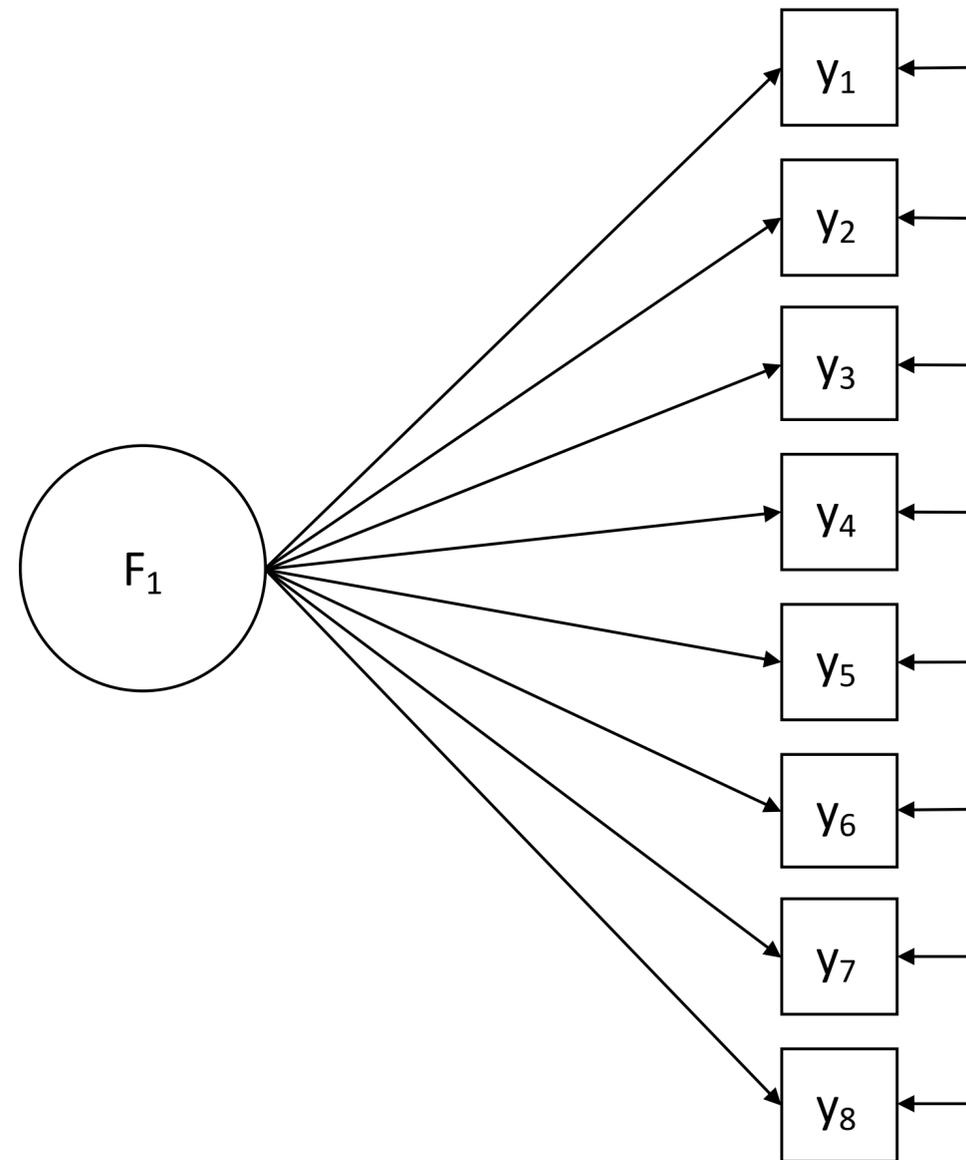
Stata. (2013) Learning the language: Path diagrams and command language. In *Stata Structural Equation Modeling Reference Manual*. College Station, Texas: Stata Press

Tu, Y. K. (2012). Directed Acyclic Graphs and Structural Equation Modelling. In *Modern Methods for Epidemiology* (pp. 191-203). Springer Netherlands.

Ahora contamos con un marco común para expresar modelos de variables latentes.  
Veamos algunos ejemplos prototípicos.

# Diagramas de variables latentes

## Ejemplo 1: Modelo de un factor

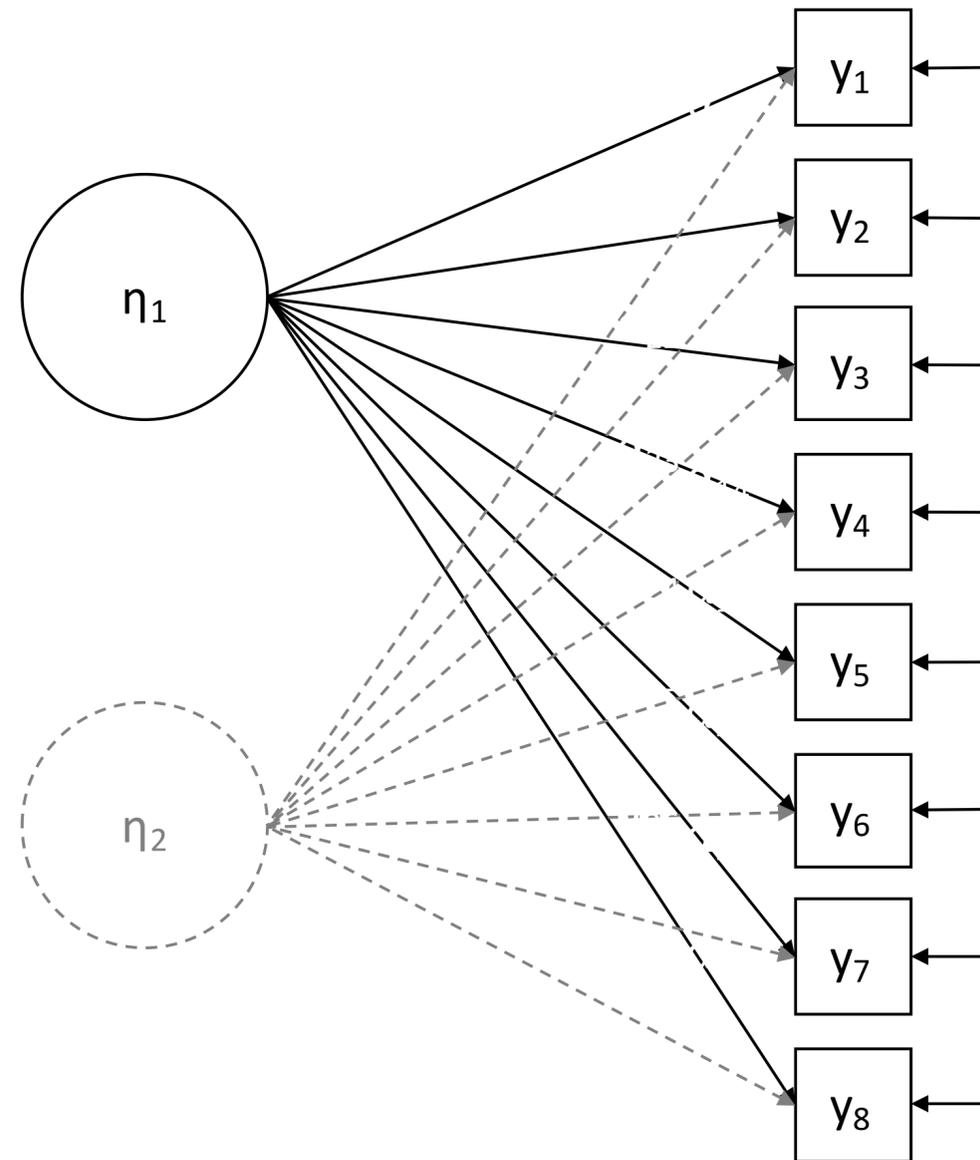


Este es un modelo de un factor, con ocho indicadores.

Los ocho indicadores  $Y$  son modelados como producto del atributo latente  $F_1$  junto con un efecto residual específico a cada indicador.

# Diagramas de variables latentes

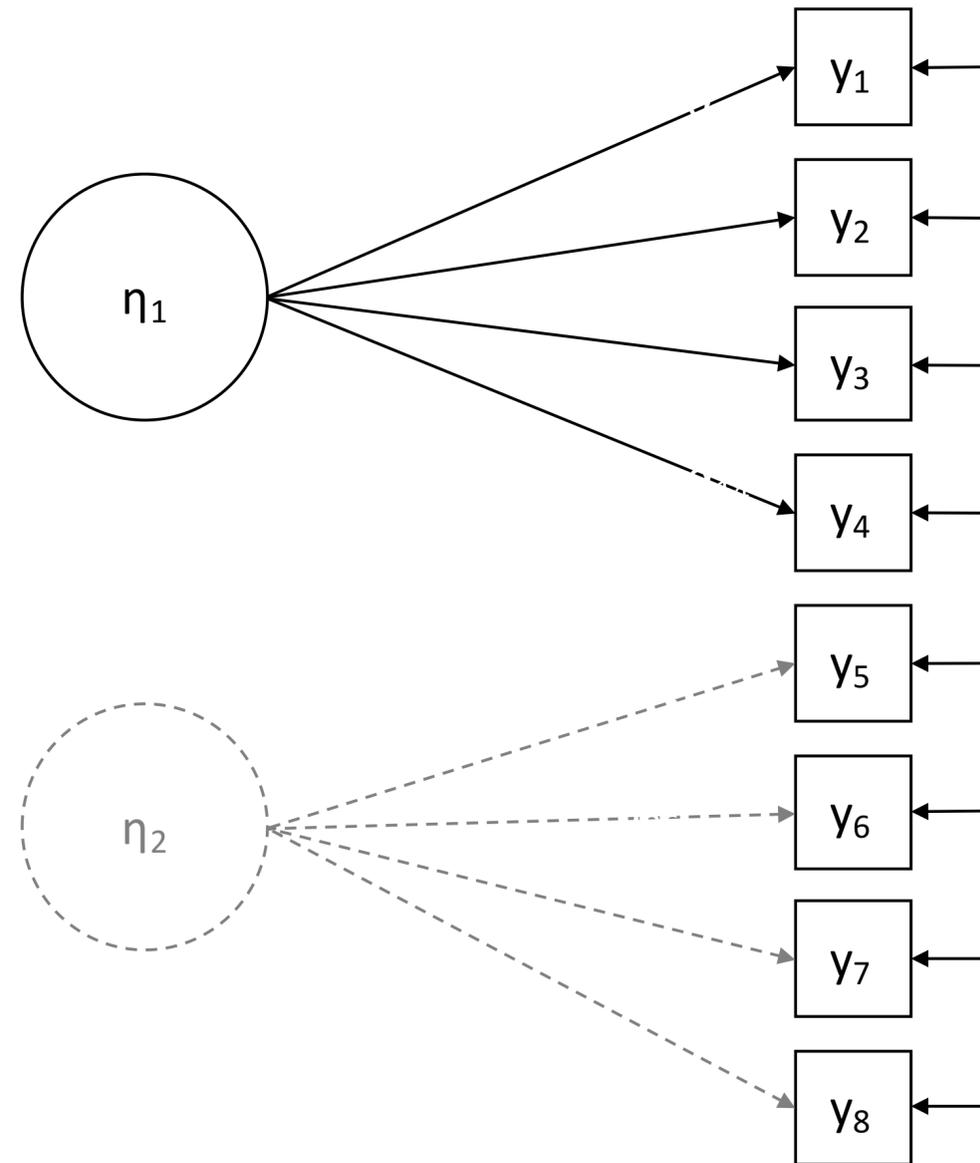
Ejemplo 2: Modelo con dos factores, análisis factorial exploratorio



Este es un modelo con dos factores, los cuales explican a ocho indicadores.

# Diagramas de variables latentes

Ejemplo 2: Modelo con dos factores, análisis factorial exploratorio



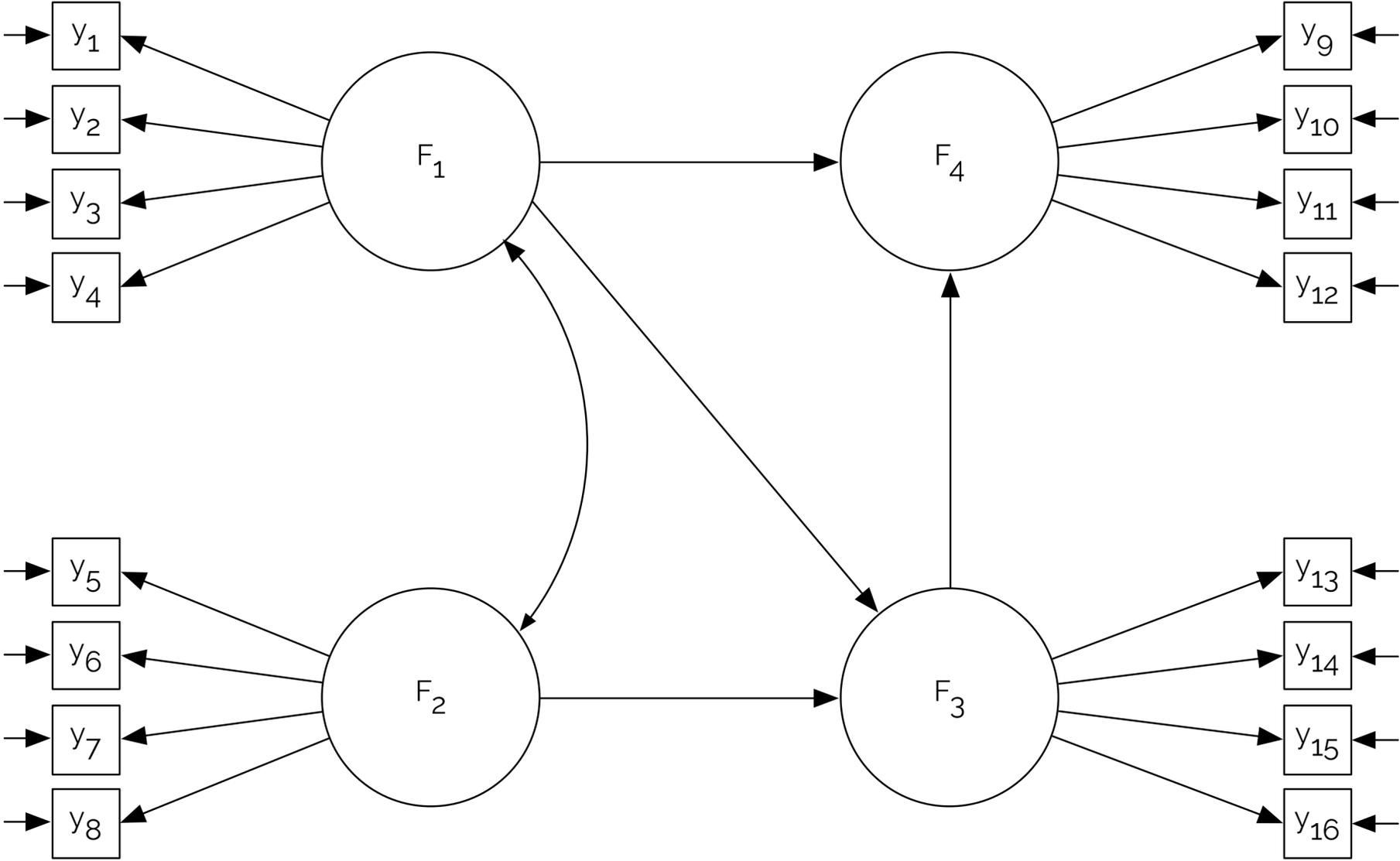
Este es un modelo con dos factores, cada uno de los cuales explica indicadores diferentes.

Empleando los componentes anteriores  
podemos especificar modelos más  
complejos.

Los siguiente son ejemplos.

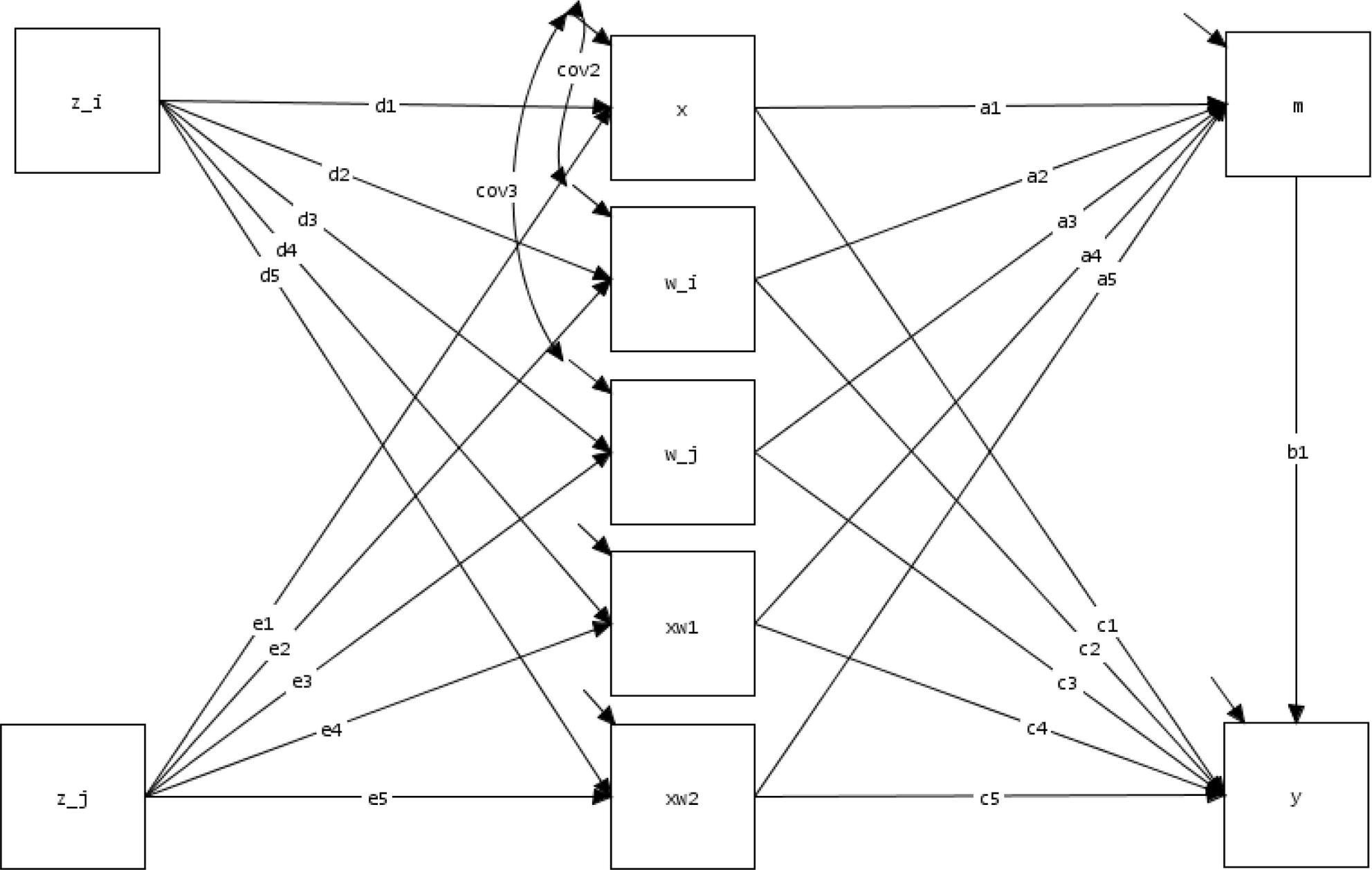
# Diagramas de variables latentes

Ejemplo 4: Modelo de ecuaciones estructurales con cuatro variables latentes y sus relaciones



# Diagramas de variables latentes

Ejemplo 5: Modelo de variables observadas, efectos indirectos y moderaciones.

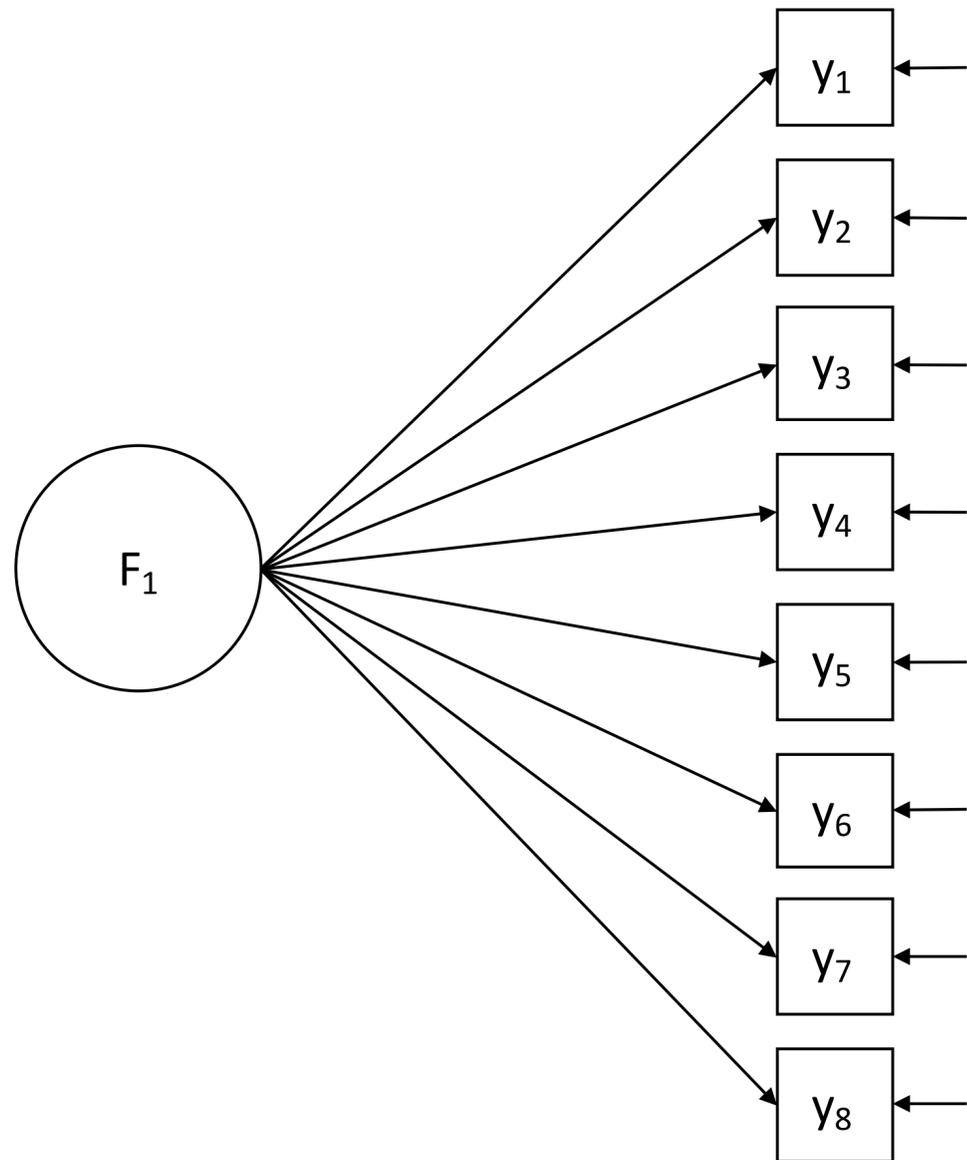


Este es un modelo más complejo, solo con variables observadas.

**Volvamos a nuestro ejemplo inicial:**  
**La estimación de una variable latente.**

# Modelo de una variable latente

Modelo de un factor: modelo tradicional

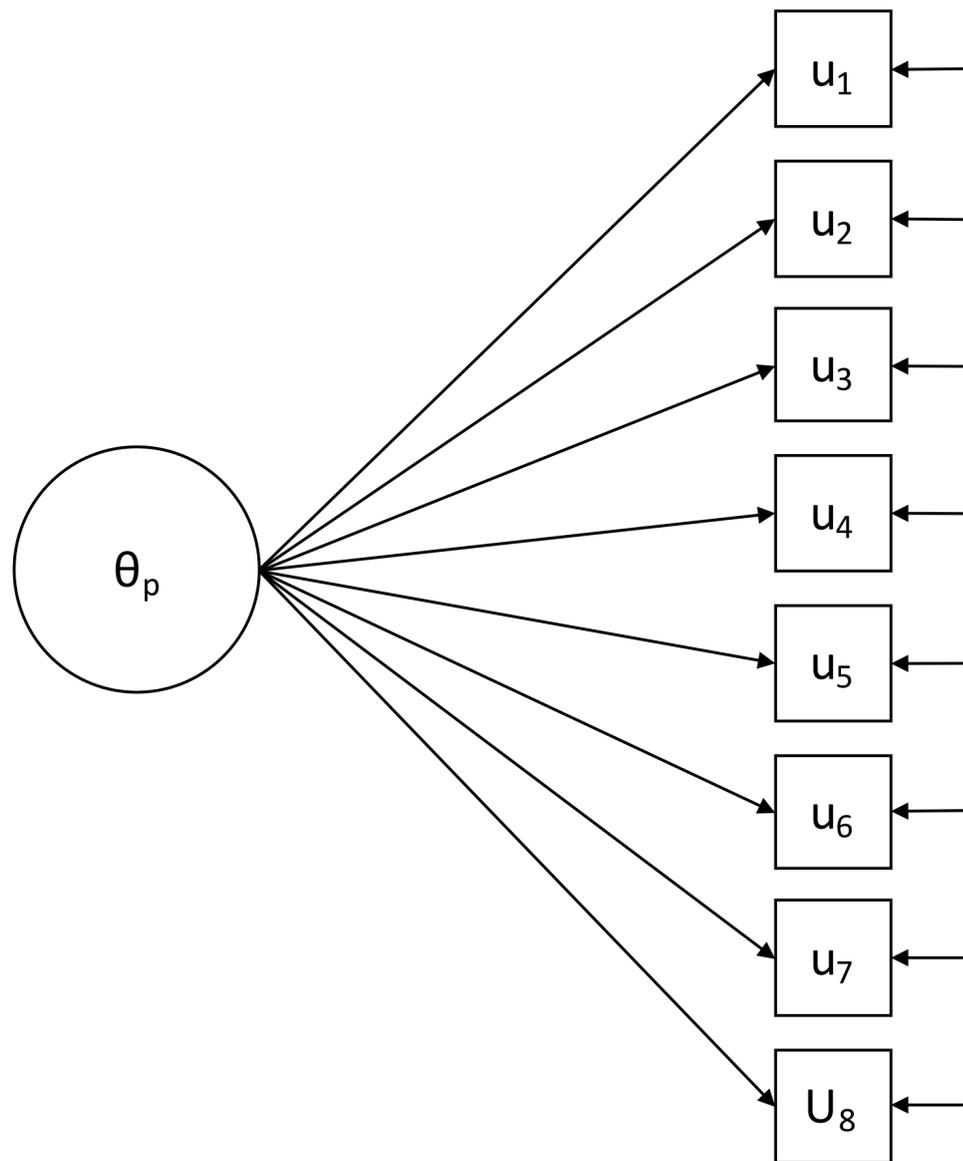


El modelo tradicional de un solo factor, en general, supone una variable latente continua ajustada sobre indicadores de estructura cuantitativa.

Sin embargo, en la versión generalizada de variables latentes, los indicadores y las variables latentes pueden tomar cualquier estructura, pudiendo ser cuantitativas, categóricas o alguna de sus variantes intermedias.

# Modelo de una variable latente

Modelo de un factor: modelo con indicadores dicotómicos



Los modelos de factor común asumen que los indicadores son cuantitativos.

En casos donde tenemos indicadores dicotómicos (e.g. respuestas correctas o incorrectas) también es posible suponer un factor. Estos modelos son conocidos como modelos de variables latentes generalizados y también como modelos de teoría de respuesta al ítem.

Podemos ver que el diagrama para modelos de factor común y modelos de teoría de respuesta al ítem son básicamente iguales, la diferencia fundamental es el tipo de indicador que se está utilizando.

Esta diferencia implica una serie de modificaciones en la representación matemática del modelo, pero en términos de la estructura del diagrama, son equivalentes.

# Modelos de variables latentes para respuestas dicotómicas: El modelo de Rasch



¿En qué consiste un modelo Rasch?  
Ilustremos sus principios con un ejemplo

# El modelo de Rasch

## Una breve introducción

El modelo de Rasch, uno de los modelos más simples en la familia de modelos IRT, indica que:

La probabilidad de una *persona*  $p$  de responder afirmativamente un *ítem*  $i$  ( $x_{ip} = 1$ ) es una función del nivel del atributo latente de la *persona*  $p$  ( $\theta_p$ ) y la dificultad del ítem  $i$  ( $\delta_i$ ):

$$\text{Probabilidad } (x_{ip} = 1) = \text{función}(\theta_p - \delta_i)$$

En términos generales, mientras más grande sea el nivel de la persona en el atributo y mientras menor sea la dificultad del ítem, más probable es que sea respondida afirmativamente/correctamente.



de Ayala, R. J. (2009). The theory and practice of ítem response theory. New York: Guilford Press.

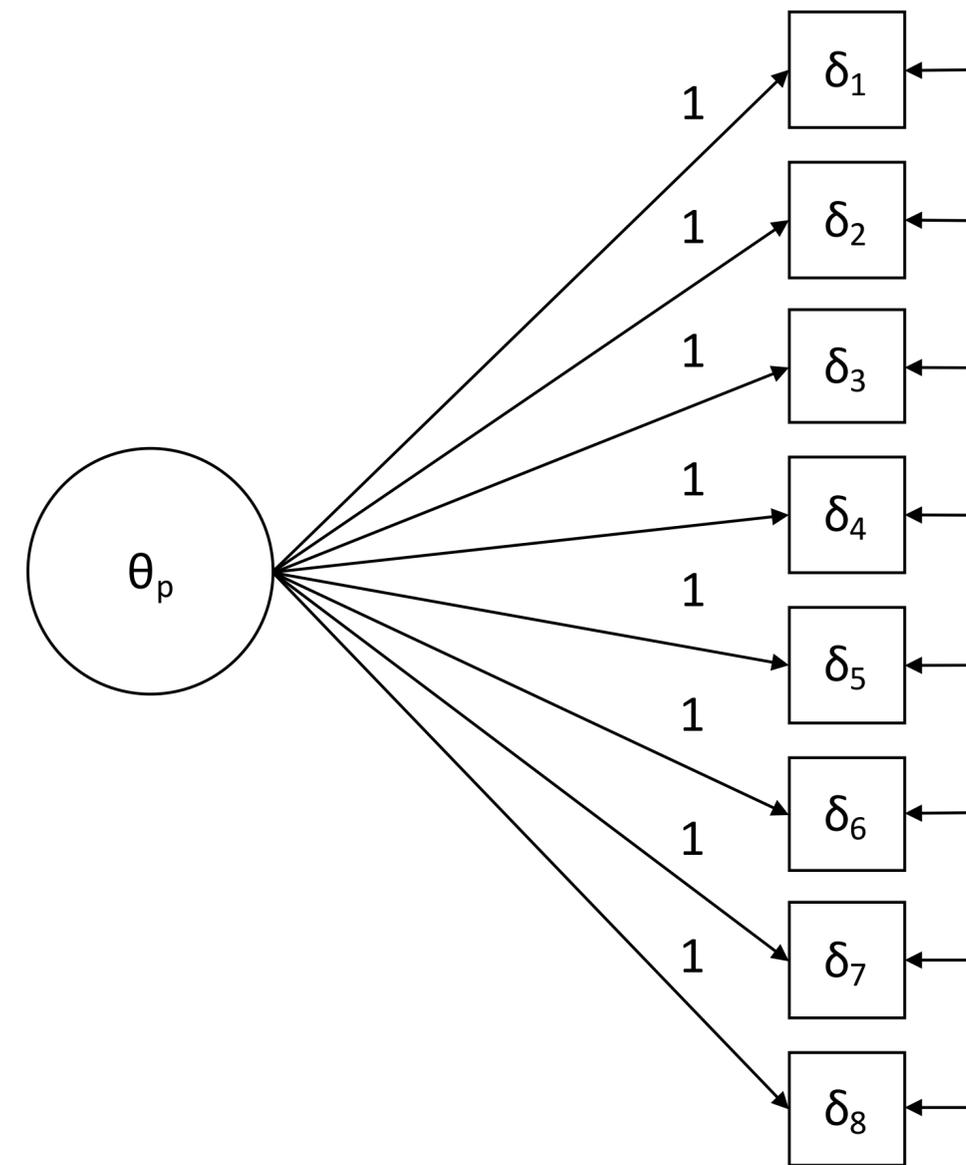
Wilson, M. (2005). Constructing measures: an ítem response modeling approach. Mahwah, N.J.: Lawrence Erlbaum Associates.

# Modelo de una variable latente

## Modelo de un factor: IRT 1PL, o Rasch

La ecuación central del modelo de Rasch muestra que la probabilidad de responder afirmativa o correctamente a una pregunta depende la diferencia entre un efecto de la persona  $\theta$  (conocimiento cívico por ejemplo) y un efecto del ítem  $\delta$  (dificultad de las preguntas de educación cívica).

Desde los modelos de variables latentes, el modelo de Rasch es comúnmente expresado en términos de un atributo latente  $\theta$  asociado a las personas (e.g. nivel de conocimiento, habilidad, desempeño, etc.) e indicadores dicotómicos sobre los que se estima su dificultad  $\delta$ .



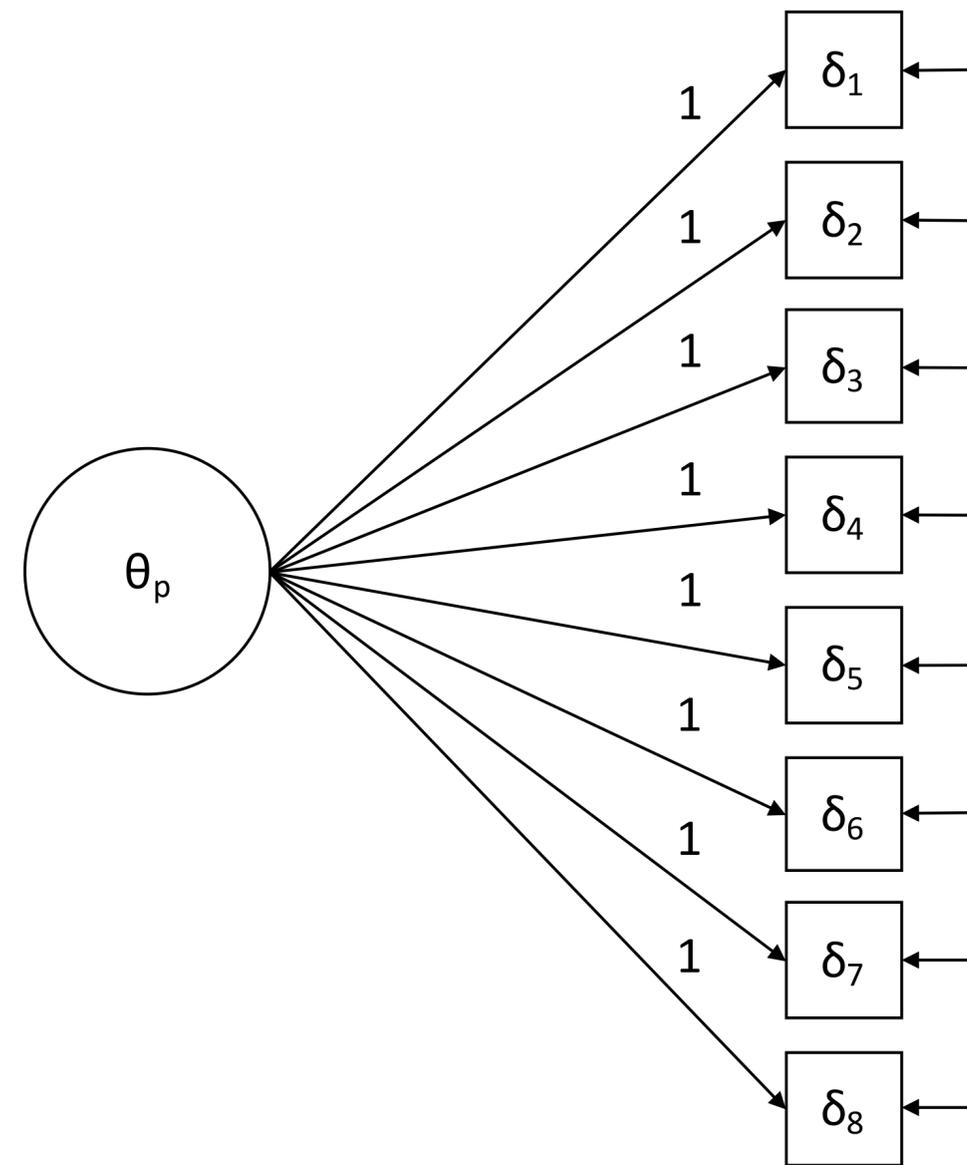
# Modelo de una variable latente

## Modelo de un factor: IRT 1PL, o Rasch

En el modelo de Rasch (y en los modelos de IRT y variables latentes en general) el error de medición puede variar de persona a persona.

Esto es una diferencia importante con el modelo de puntaje verdadero, donde el error de medición es igual para todos quienes responden una misma prueba.

Los modelos de variables latentes estiman un error de medición diferente dependiendo de la ubicación en la que se estima el efecto de las personas y la cantidad de información disponible (e.g. qué preguntas respondió cada persona).

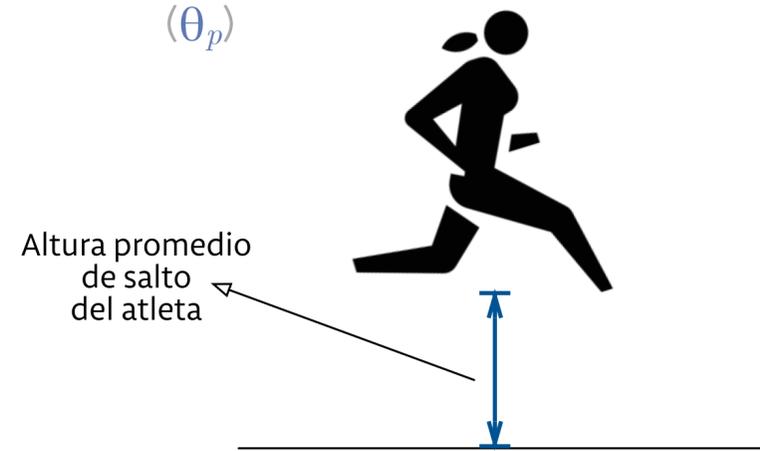


# Modelos Rasch

## Introducción

Nivel de la persona

$(\theta_p)$

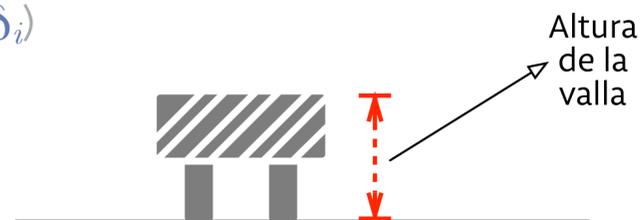


En un modelo Rasch, la habilidad y la dificultad son expresados en una misma escala.

Ilustremos esto con el salto de vallas. En este ejemplo **altura** es el atributo de interés.

Nivel del ítem

$(\delta_i)$



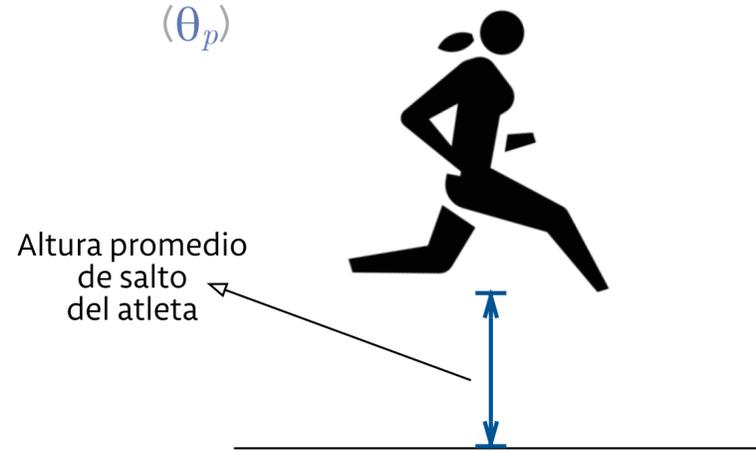
En un modelo de Rasch (como en otros modelos de variables latentes) este atributo sería no observable, pero en este ejemplo ilustramos las ideas centrales con una variable observada.

# Modelos Rasch

## Introducción

Nivel de la persona

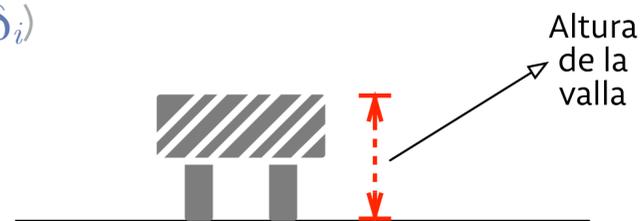
$(\theta_p)$



Altura promedio de salto del atleta

Nivel del ítem

$(\delta_i)$



Altura de la valla

Recordemos que:

Probabilidad ( $x_{ip} = 1$ ) = función( $\theta_p - \delta_i$ )

Esta ecuación muestra que elemento central del modelo es la diferencia entre el parámetro que representa el nivel del atributo en la persona  $\theta_p$  y el parámetro que representa el nivel de atributo en el ítem  $\delta_i$ .

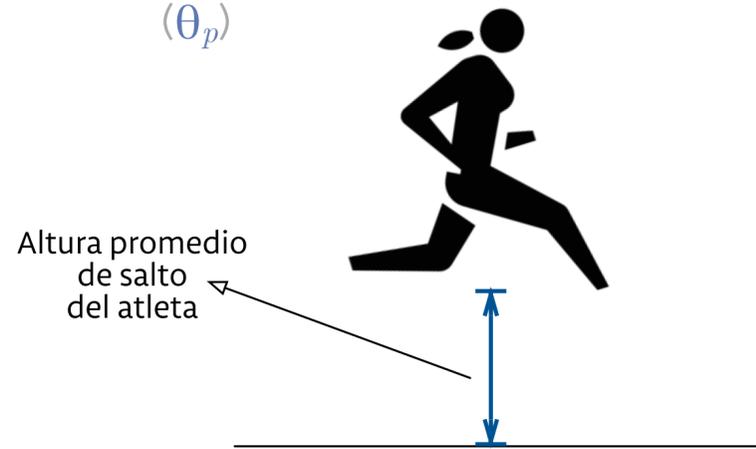
En nuestro ejemplo de vallas, la altura promedio del salto del atleta es representada por  $\theta_p$  y la altura de la valla es representada por  $\delta_i$ .

# Modelos Rasch

## Introducción

Nivel de la persona

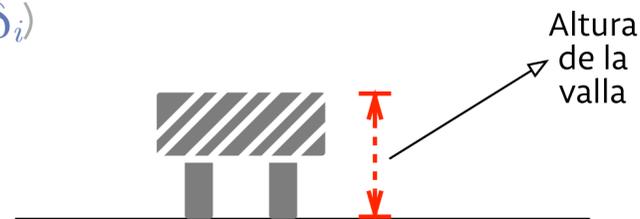
$(\theta_p)$



Altura promedio de salto del atleta

Nivel del ítem

$(\delta_i)$



Altura de la valla

La intuición entonces es que la probabilidad de saltar correctamente aumenta en la medida que:

$$\theta_p > \delta_i$$

Nivel de la persona > Nivel del ítem

De modo que la probabilidad de tener éxito saltando:

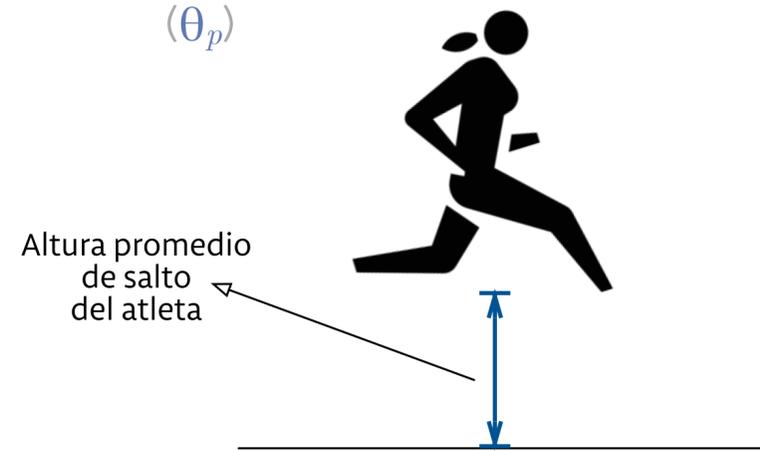
- aumenta en la medida que  $\theta_p - \delta_i$  sea mayor que cero y aumente.
- disminuye en la medida que  $\theta_p - \delta_i$  sea menor que cero y disminuya.
- cuando  $\theta_p - \delta_i$  es igual a cero (cuando la persona y valla tiene el mismo nivel) la probabilidad de saltar exitosamente es igual a la probabilidad de fallar en el salto.

# Modelos Rasch

## Introducción

Nivel de la persona

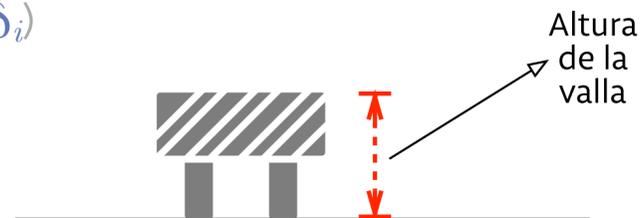
$(\theta_p)$



En este ejemplo se puede ver que tanto el nivel de la persona como el nivel de la valla son modelados en una misma escala: altura.

Nivel del ítem

$(\delta_i)$



Esto es también el caso en el modelo de Rasch (y los modelos de IRT en general), de forma que las personas e ítems son representados en una escala común.

# Modelos Rasch

## Introducción

*En una prueba de conocimiento de matemática, la persona tiene un nivel de conocimiento y el ítem tiene también un nivel de conocimiento.*

*El nivel de la persona se interpreta como el conocimiento que la persona posee, y el nivel del ítem como el conocimiento que este demanda para ser respondido correctamente.*

*De forma similar al ejemplo de las vallas, creemos que una persona tiene más probabilidad de responder correctamente cuando su nivel de conocimiento supera el nivel de conocimiento que demanda la pregunta.*

# Modelos Rasch

## Introducción

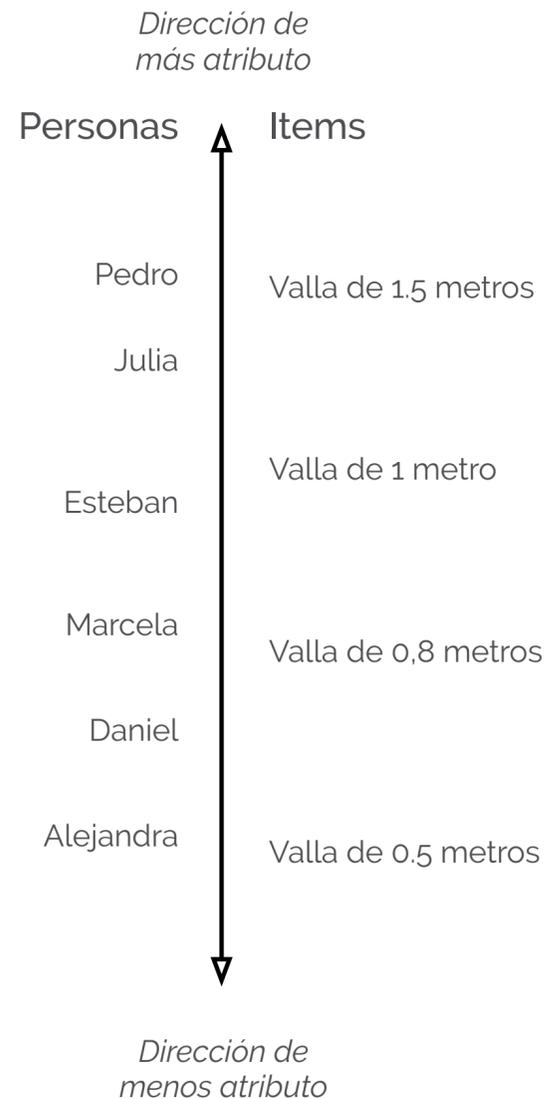
*En un cuestionario de apoyo al gobierno, la persona tiene un nivel de apoyo y el ítem tiene también un nivel de apoyo.*

*El nivel de la persona se interpreta como el apoyo al gobierno que la persona presenta, y el nivel del ítem como el nivel de apoyo al gobierno que se requiere para estar de acuerdo con esa afirmación.*

*De forma similar al ejemplo de las vallas, creemos que una persona tiene más probabilidad de estar de acuerdo con una afirmación cuando su nivel de apoyo supera el nivel de apoyo que demanda la afirmación.*

# Modelos Rasch

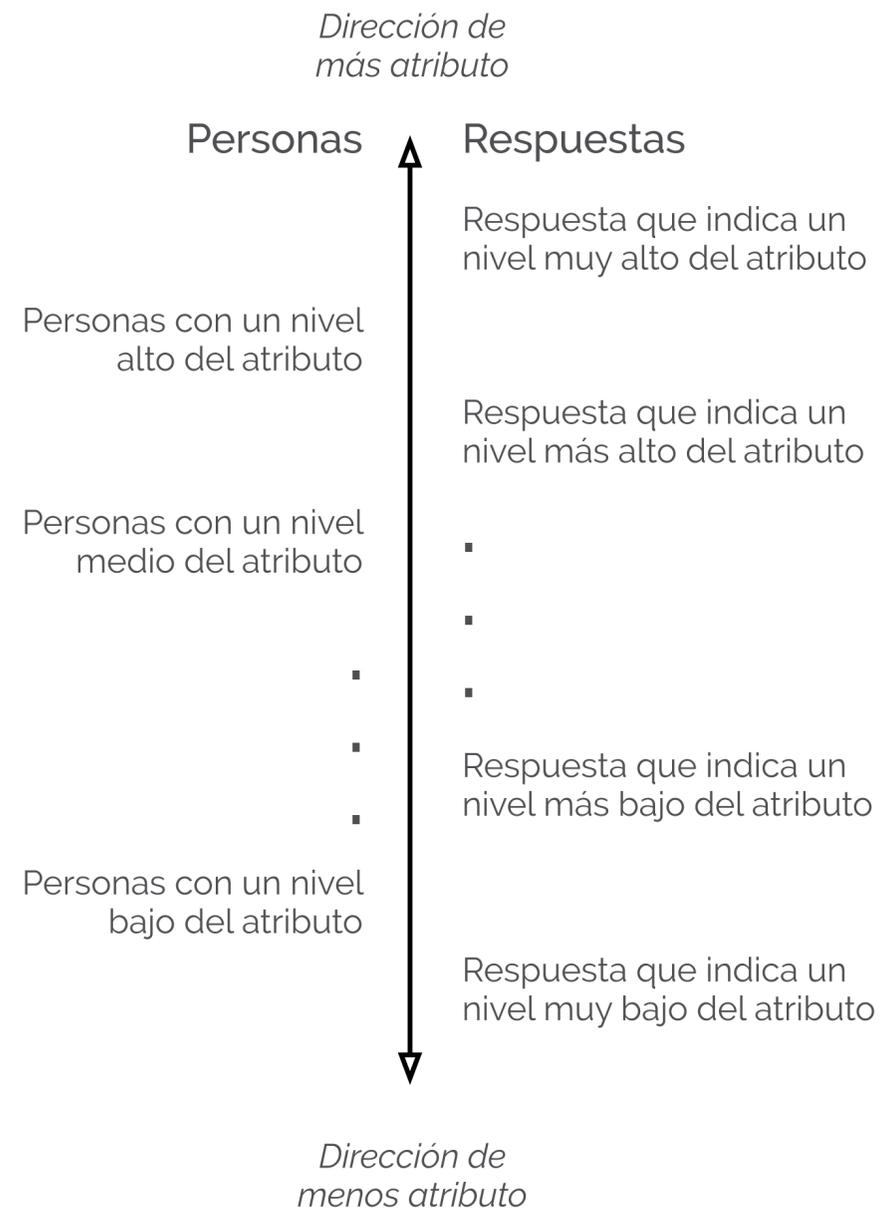
## Introducción



Cuando se analizan las respuestas de personas a una prueba el modelo de Rasch nos permite crear una versión empírica del mapa de constructo, posicionando a cada persona y a cada ítem en el atributo latente.

# Modelos Rasch

## Introducción



En el caso de los modelos Rasch, la estrategia de selección de ítems está guiada por el nivel del atributo a medir sobre el que se quieren hacer inferencias.

La idea es mapear los diferentes niveles del atributo que nos interesan. No solo tener ítems de alta, baja o moderada dificultad, sino cubrir el espectro de habilidad que nos interesa.

Mientras menos ítems tenemos en un sector del atributo latente, menos información tenemos sobre esa sección del atributo latente.



de Ayala, R. J. (2009). The theory and practice of item response theory. New York: Guilford Press.

Wilson, M. (2005). Constructing measures: an item response modeling approach. Mahwah, N.J.: Lawrence Erlbaum Associates.

# Modelos Rasch

## Supuestos

**El uso de modelos de variables latentes en general requiere que se cumplan varios supuestos. La idea fundamental es que si no se cumplen los supuestos de un modelo, no es razonable utilizarlo para hacer inferencias.**

De forma conceptual, esperamos que tenga sentido modelar el atributo como una cantidad, ya que al ser un atributo latente, estamos haciendo una hipótesis respecto a su estructura.

Estadísticamente, al analizar una prueba usando el modelo de variables latentes esperamos se cumplan a lo menos dos supuestos fundamentales:

1. Que las respuestas a los ítems sean condicionalmente independientes a la variable latente, lo que significa que toda correlación entre las respuestas a los ítems (variable observadas) es producto de la variable latente.
2. Que las respuestas a los ítems de la prueba son producto de una sola dimensión. Esto es expresado tradicionalmente diciendo que estos modelos suponen que “solo estamos midiendo un solo atributo”.

Existen modelos más complejos que pueden relajar la necesidad de hacer estos supuestos (e.g. modelos multidimensionales), pero un análisis de una prueba con modelo de Rasch simple descansa necesariamente sobre estos supuestos.

**Al utilizar modelos de variables latentes es posible generar múltiples indicadores y estadísticos de ajuste que deben ser evaluados para determinar si es razonable utilizar dichos modelos.**



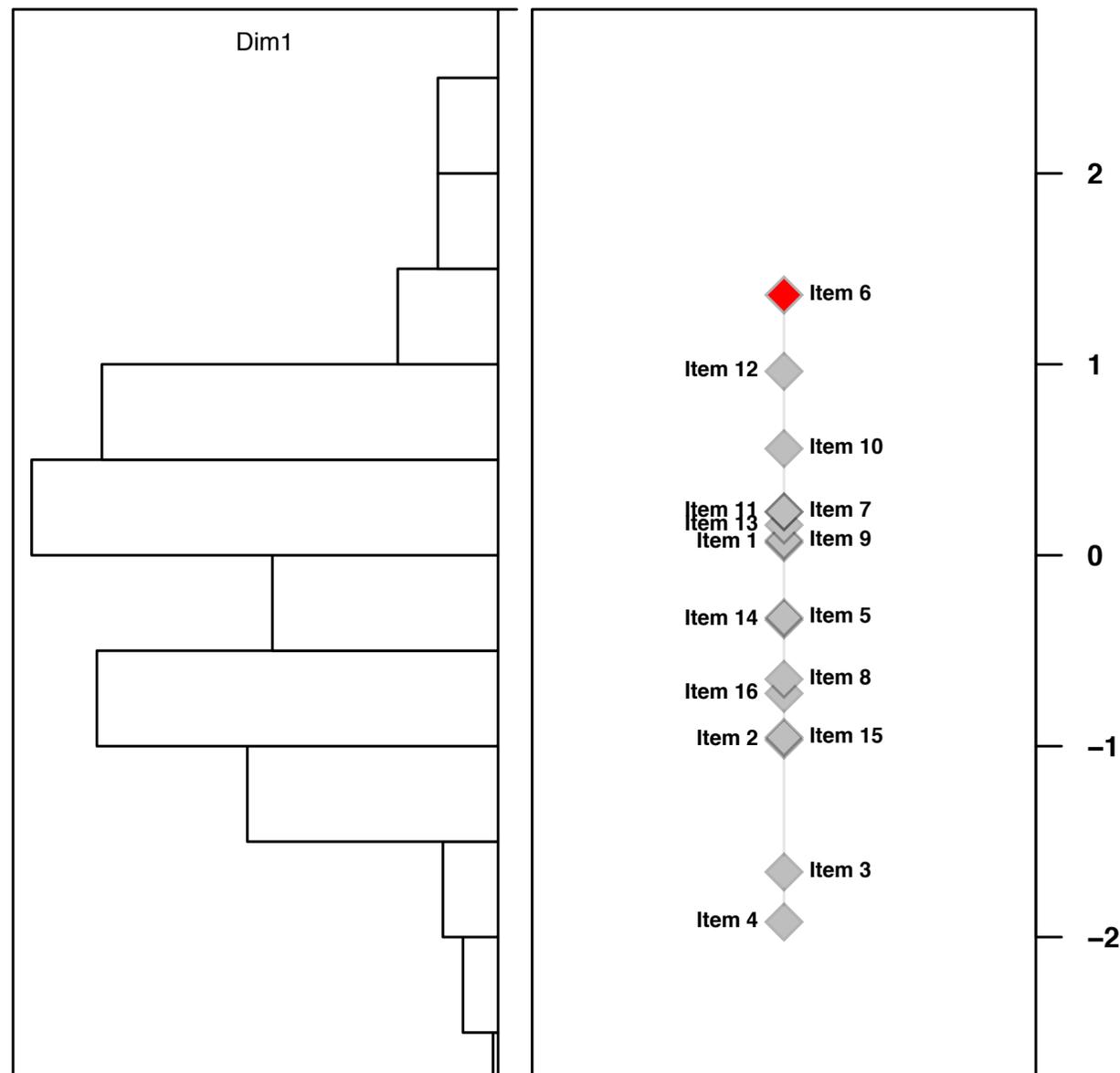
de Ayala, R. J. (2009). The theory and practice of item response theory. New York: Guilford Press.

Wilson, M. (2005). Constructing measures: an item response modeling approach. Mahwah, N.J.: Lawrence Erlbaum Associates.

¿Cómo se ordenan los ítems de cívica?  
Mapa de constructo empírico de los ítems  
ICCS 2009, módulo latinoamericano

# Modelos Rasch

## Introducción

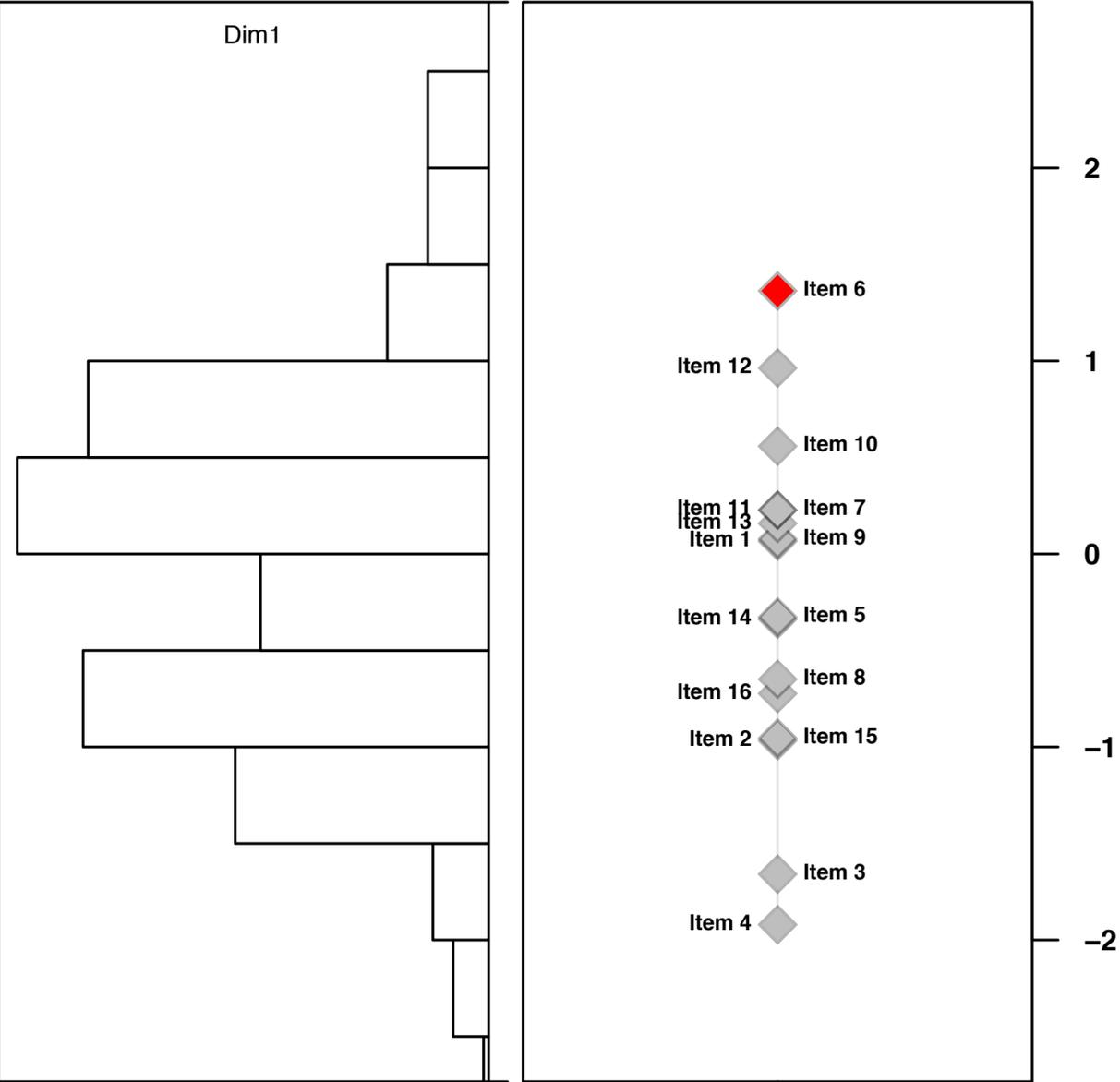


Observando el mapa de ítems-personas, o mapa de constructo empírico, podemos ver que el ítem 6 es informativo de alta habilidad.

Si lo elimináramos, perderíamos información en ese sector de la variable latente.

# Modelos Rasch

## Introducción



¿Cuál de las siguientes fue una de las consecuencias de las dictaduras latinoamericanas del siglo pasado?

Un líder que se comporta democráticamente

¿Qué caracteriza a un gobierno autoritario?

¿Cuál de las siguientes es una razón por la que existen las leyes de tránsito?

# Marcos generales de modelos de variables latentes



# Modelos de variables latentes

## Modelos estadísticos

Los modelos de variables latentes surgen con Spearman a comienzos del siglo 20.

No obstante, la adopción de marcos estadísticos generales que incorporan múltiples tipos de modelos de variables latentes como casos especiales ha aumentado considerablemente en las últimas décadas (1990-2010).

Estos marcos estadísticos generales presentan familias de modelos estadísticos que nos permiten formalizar y estimar relaciones entre variables. El desarrollo de estos marcos ha ido de la mano con el desarrollo de programas estadísticos que permiten estimar cualquier modelo que pueda ser formulado como un caso especial dentro de una familia general de modelos.



# Modelos de variables latentes

## Modelos estadísticos



*[...] La omnipresencia de las variables latentes no es comúnmente reconocida, quizás porque éstas reciben diferentes nombres en diferentes literaturas, tales como efectos aleatorios, factores comunes y clases latentes.*

— Skrandal & Rabe-Hesketh 2004

El marco general de variables latentes nos permite formalizar y expresar mediante modelos las relaciones entre variables observadas y no observadas. Estos modelos pueden incluir, por ejemplo, a los modelos multinivel y los modelos de teoría de respuesta al ítem (IRT).

El marco general de variables latentes nos permite pensar especificaciones anteriores, como si fueran casos especiales de un mismo marco.

# Modelos de variables latentes

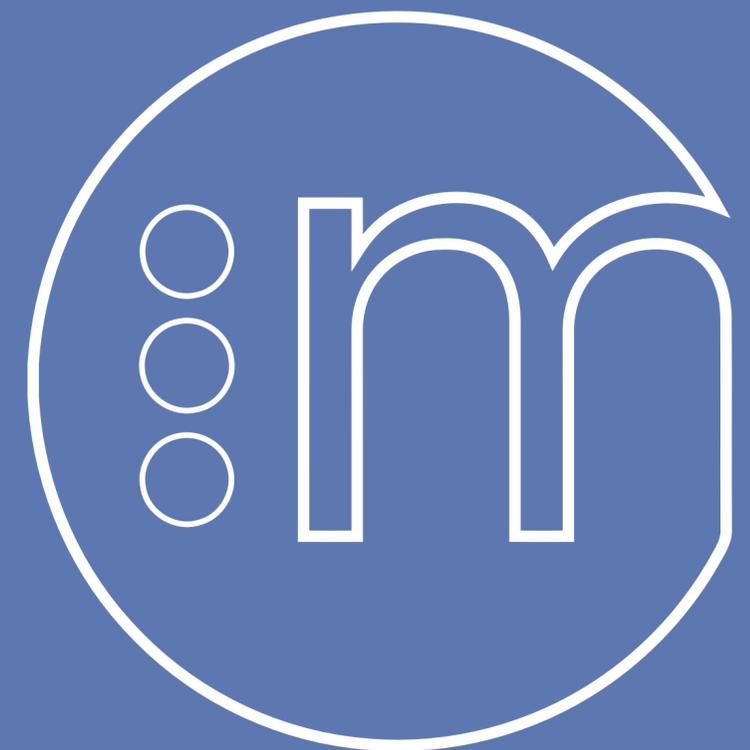
## Modelos estadísticos

Ahora que compartimos este marco de referencia, podemos volver sobre algunas de las preguntas que hemos visto en clases anteriores, tales como ejemplos de sesgo, y los modelos asociados a evaluar la dimensionalidad de un conjunto de ítems.

El uso de modelos de variables latentes para formular estas temáticas nos permite examinarlos empíricamente en base a las respuestas recogidas. Esto nos da una perspectiva adicional al momento de hacer juicios respecto a los ítems.

# Problemas de Validez

...pensados desde un marco de variables latentes



¿Cómo evaluar el problema de tener varios contenidos pero un solo atributo de interés?  
Una alternativa son los modelos bifactoriales

# Aplicaciones de MVL

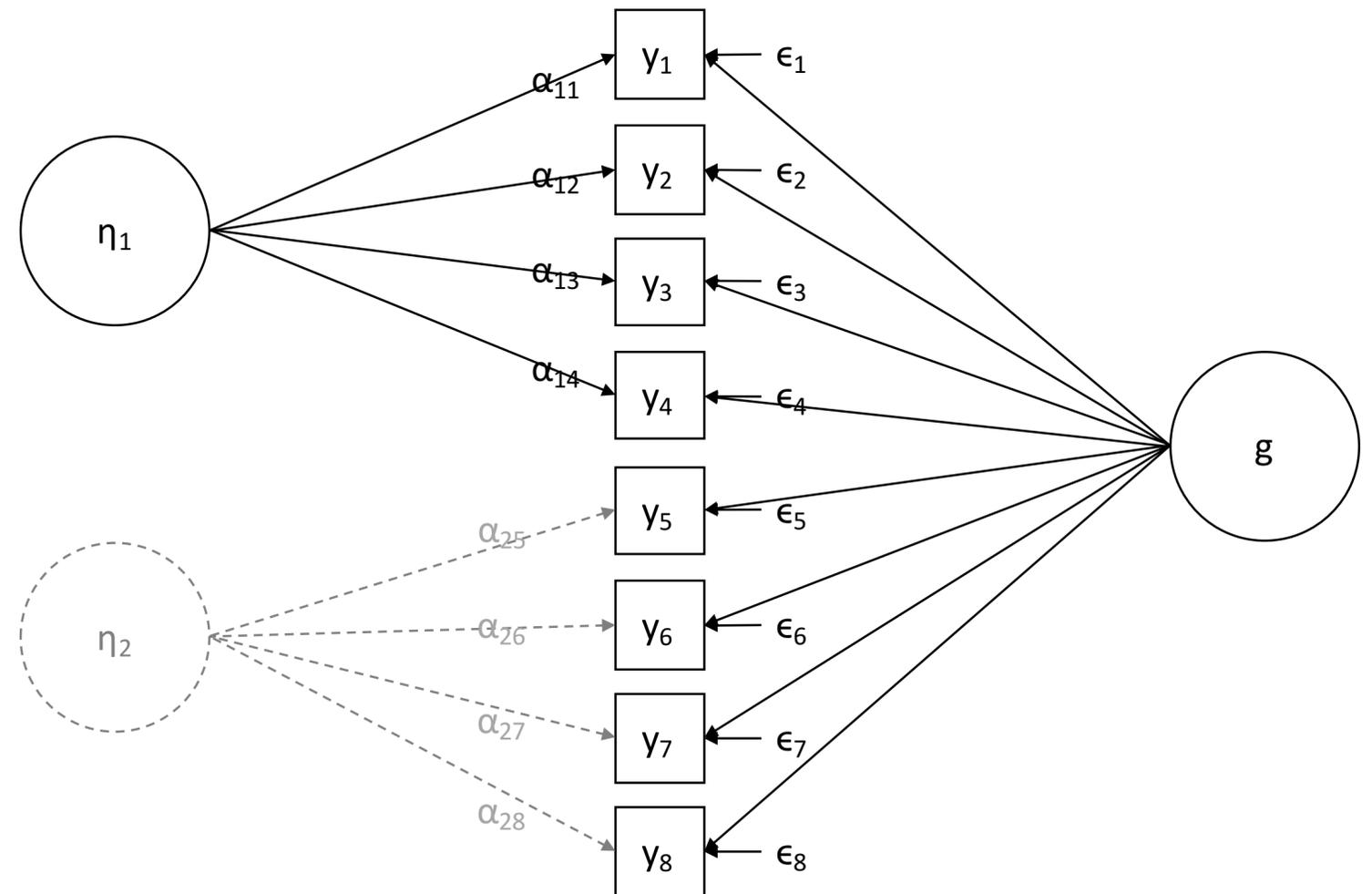
## Modelos bifactoriales

En clases anteriores hemos visto describir en más de una ocasión, la idea general de que en una prueba de matemáticas, se evalúa más de un contenido.

Por ejemplo, una prueba puede contener items de geometría y de álgebra.

¿Cómo aseguramos que tenemos una sola dimensión y no dos dimensiones?

Los modelos **teslet** y los modelos **bifactoriales** pueden ser usados con este propósito.



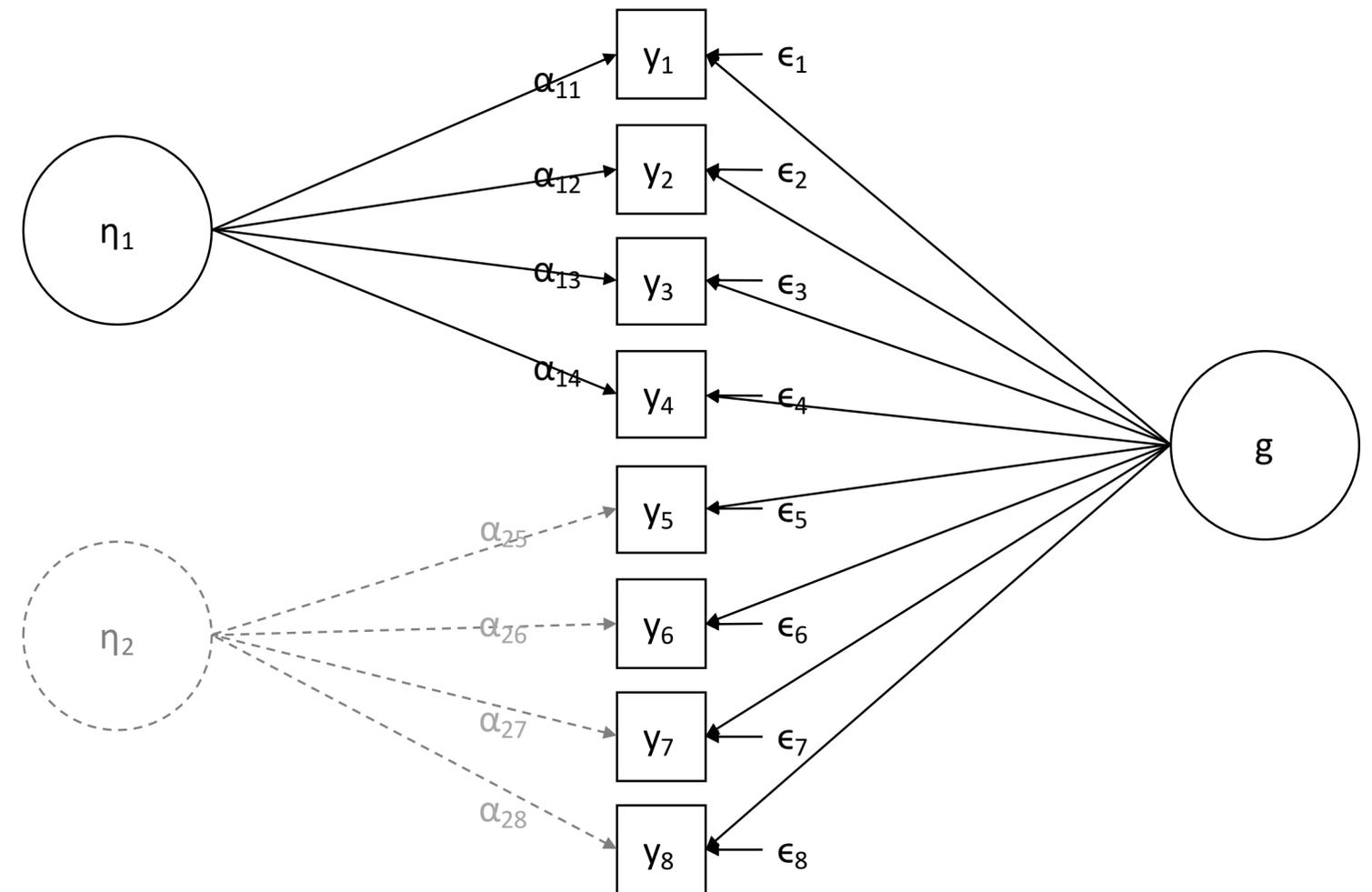
Reise, S. P., Morizot, J., & Hays, R. D. (2007). The role of the bifactor model in resolving dimensionality issues in health outcomes measures. *Quality of Life Research*, 16(SUPPL. 1), 19–31. <http://doi.org/10.1007/s11136-007-9183-7>

# Aplicaciones de MVL

## Modelos bifactoriales

Se comparan las relaciones entre el factor general (matemáticas), en contraste a las las relaciones entre los factores específicos (i.e. geometría y álgebra).

Si el supuesto de unidimensionalidad se cumple, se espera que los factores específicos posean relaciones muy pequeñas con los indicadores.



Reise, S. P., Morizot, J., & Hays, R. D. (2007). The role of the bifactor model in resolving dimensionality issues in health outcomes measures. *Quality of Life Research*, 16(SUPPL. 1), 19–31. <http://doi.org/10.1007/s11136-007-9183-7>

¿Cómo evaluar el sesgo de ítems?  
¿Cómo saber si es sólo diferencia entre grupos o problemas con los ítems?

# Aplicaciones de MVL

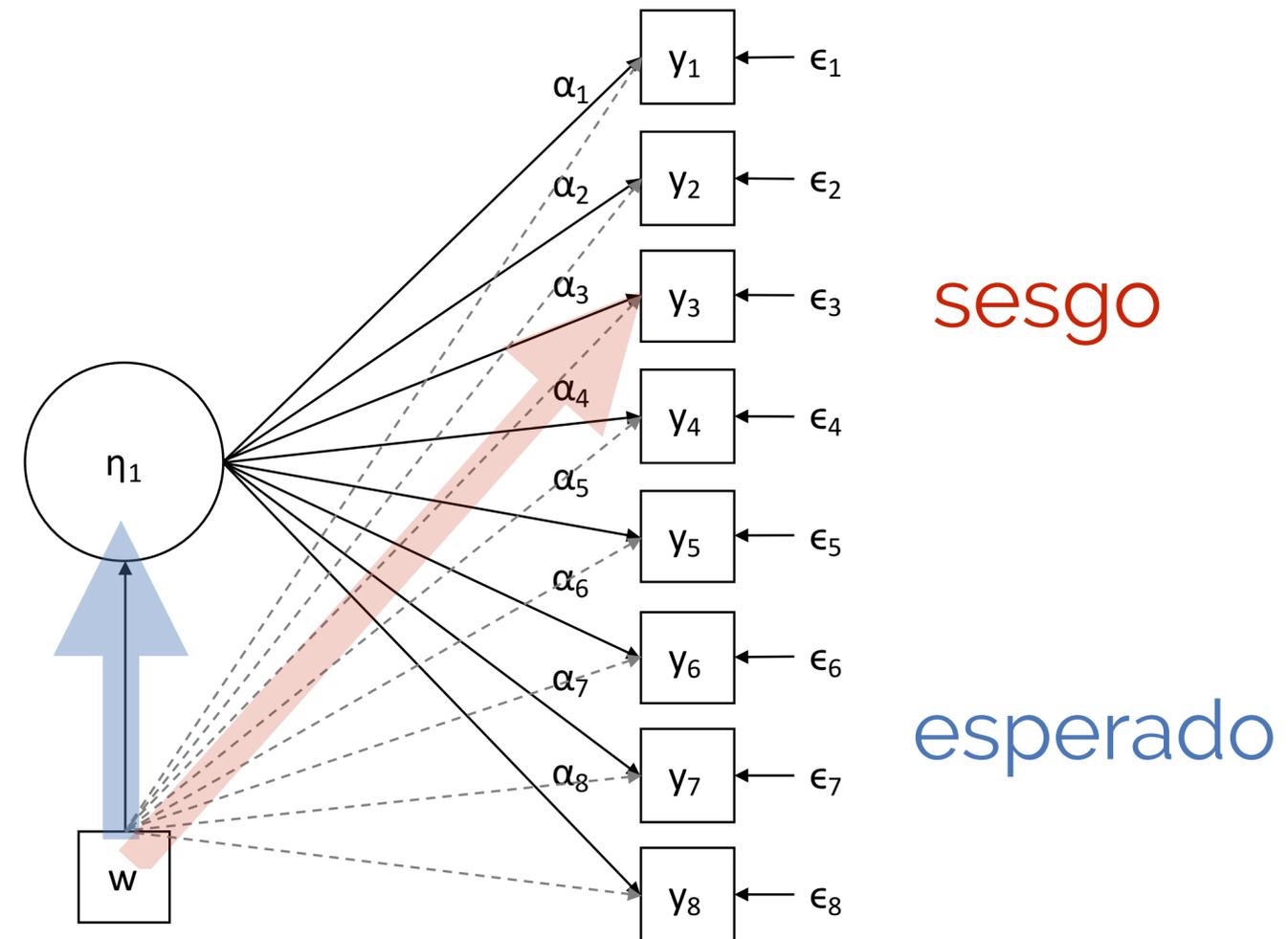
## Sesgo de ítems en SEM

Una forma de evaluar si existiera sesgo de ítems para un grupo, es especificar una covariable en el modelo que representa al grupo.

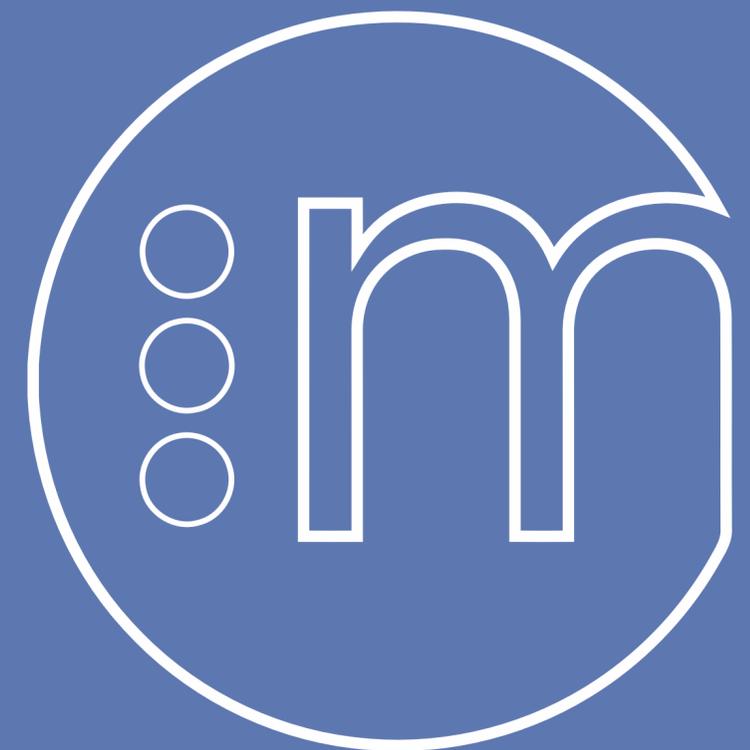
Esta ha de contar con una relación al atributo medido, y adicionalmente, una serie de relaciones a los indicadores observados.

Se espera, que en caso de ausencia de sesgo, esta covariable, que podría ser nacionalidad, sexo, nivel socioeconómico u otra; sólo se relacionara de forma directa al atributo.

En caso de sesgo, las relaciones adicionales de la covariable a los indicadores serían diferentes de cero.



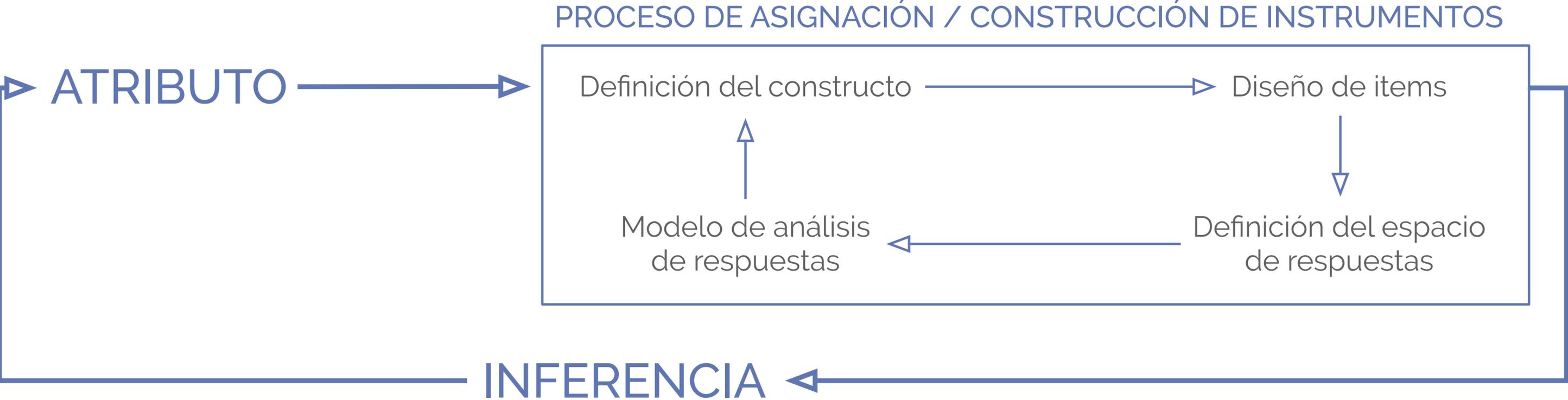
# Resumen del proceso



# Medición en ciencias sociales

Atributos, el proceso de asignación y las inferencias

Recordemos nuevamente los tres elementos involucrados en el proceso de medición en las ciencias sociales:



# Créditos

Clases preparadas por el área de investigación MIDE – agosto 2016

Diego Carrasco — María Inés Godoy — Daniela Jiménez — David Torres Iribarra

y agradecimientos a Mauricio Rivera

Todos los símbolos provienen de [thenounproject.com](http://thenounproject.com)

 Creado por Quinn Keaveney

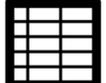
 Creado por Jonathan Gibson

 Creado por [designify.me](http://designify.me)

 Creado por Gilbert Bages

 Creado por Agniraj Chatterji

 Creado por TMD

 Creado por [useiconic.com](http://useiconic.com)

 Creado por Christopher Smith

 Creado por Hannah Strobel

 Creado por Jaime Carrion

 Creado por Gonzalo Bravo

 Creado por Takao Umehara

 Creado por Krisada, TH

 Creado por Alexander Blagochevsky

# Introducción a la medición en las ciencias sociales

Sesión 4 - Modelos de Análisis

Área de Investigación



**Centro UC**  
Medición - MIDE