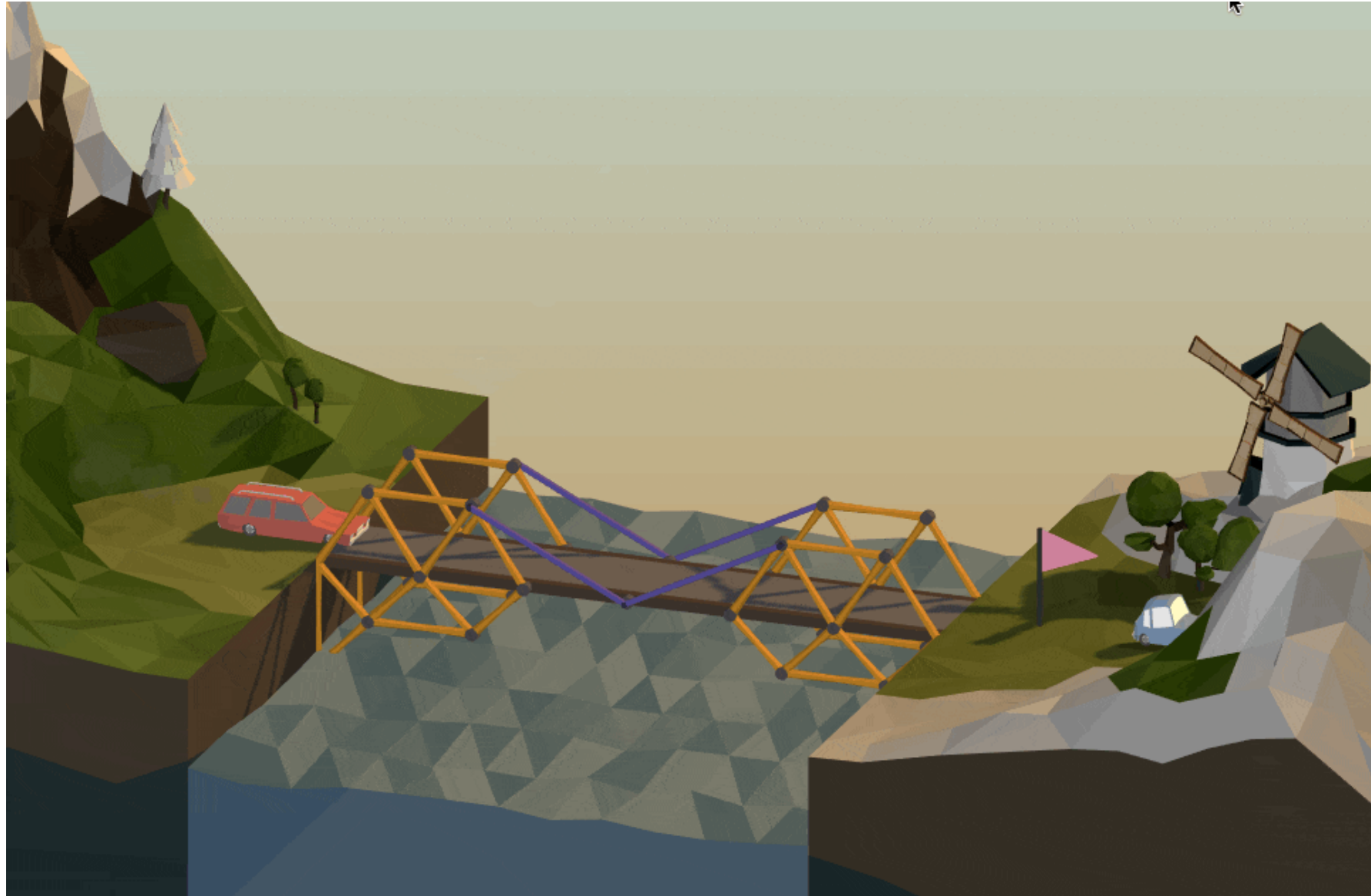


Validez en medición

Validez

- Contenido
- Cara
- Criterio
- Concorrente
- **Constructo**

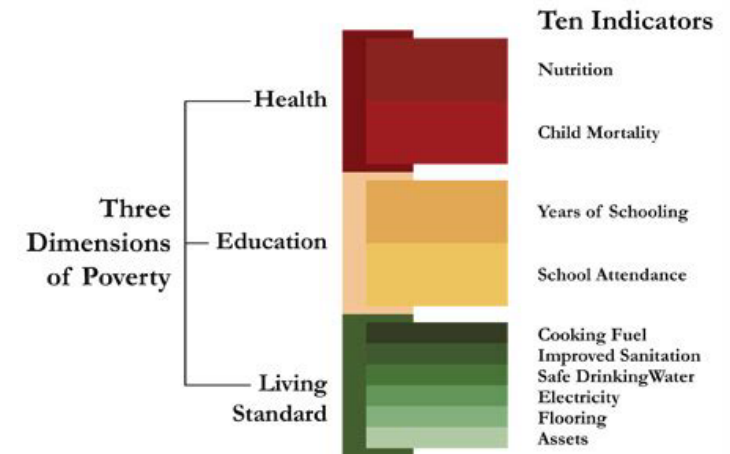


Validez de contenido / sistemática

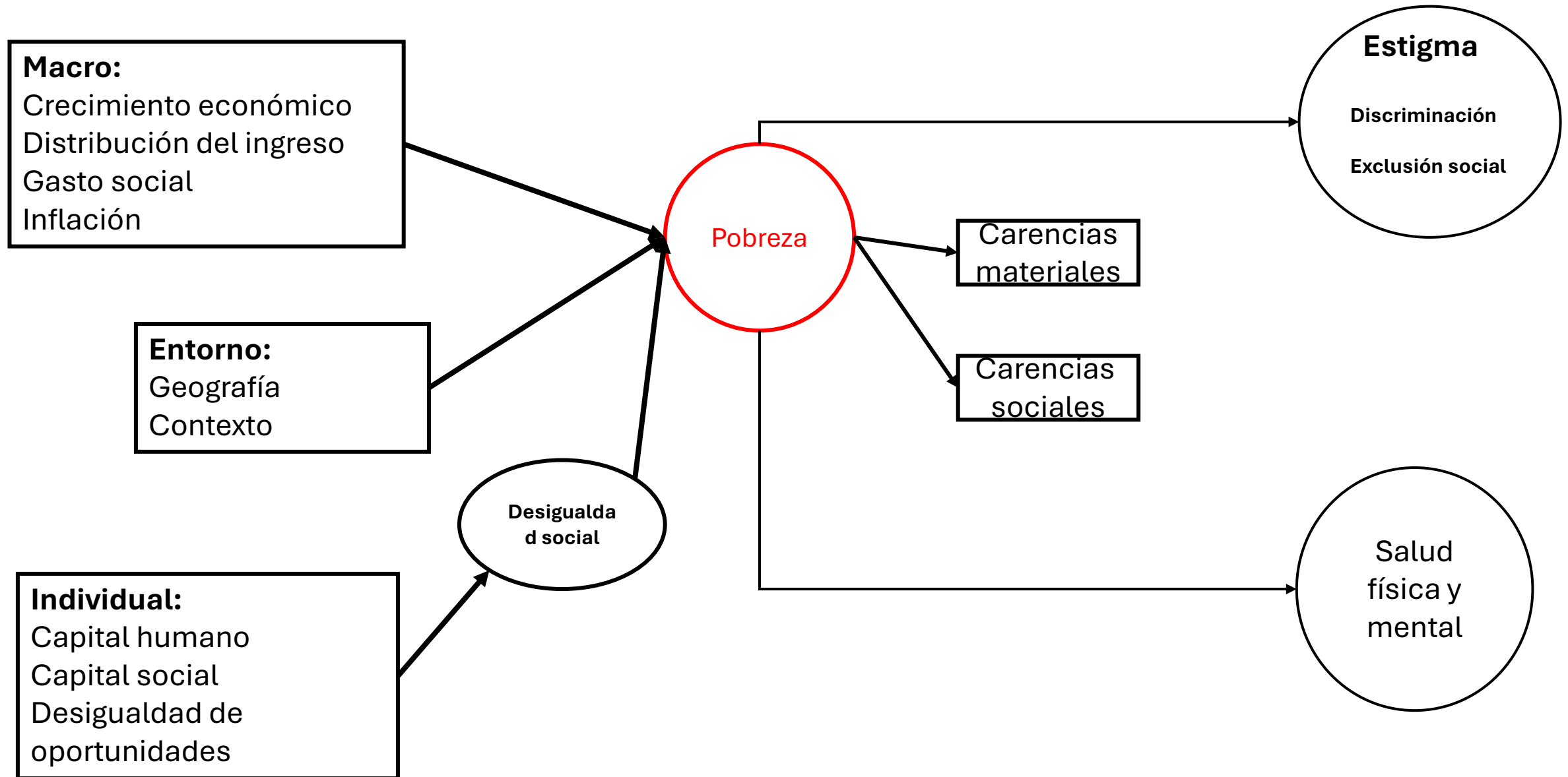
- Refiere a la validez del modelo teórico: consistencia interna del modelo teórico
- Modelo científico:
 - Definición
 - Distinción entre causas y consecuencias
- La validez de contenido busca atar los ítems a las razones teóricas detrás de su formulación
- En algunos casos se usan marcos normativos (derechos humanos) o teóricos (necesidades básicas insatisfechas) para justificar su inclusión en las escalas

Validez de contenido

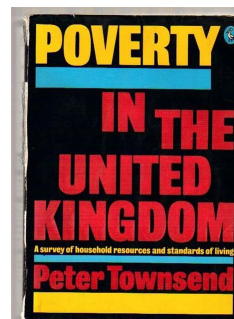
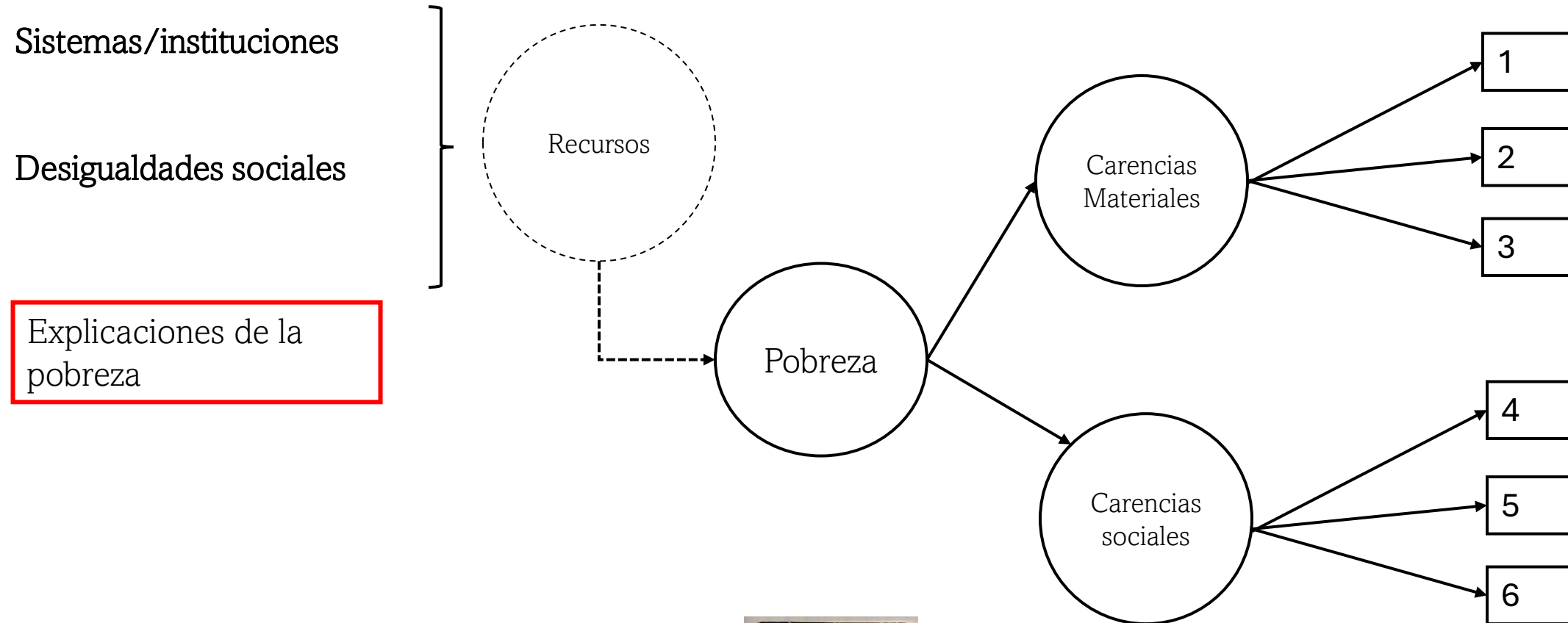
- OK! Las correlaciones entre distintos tests y algunos predictores del fenómeno puede ser útil pero...
- ¿Dónde quedó la teoría detrás de la selección y formulación de cada ítem?
- Como nos decía Goldstein, la TRI no es una teoría de las preguntas sino un modelo probabilístico sobre los 1 y 0 ´s que vemos



Explicaciones de pobreza factores

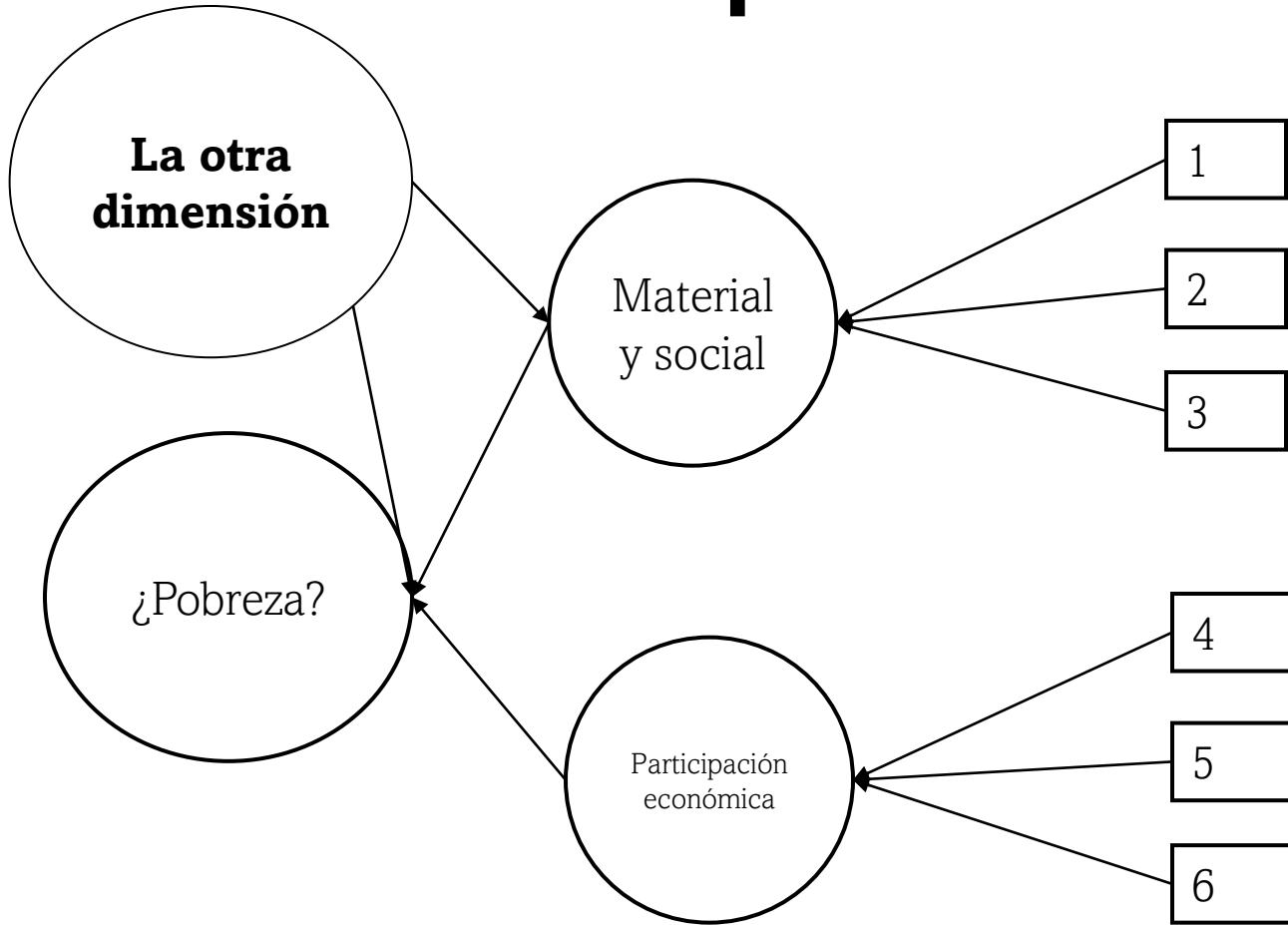


Sin modelo, no hay medición



Modelo convencional de la teoría contemporánea de la medición

Esquema de medición no científico



Modelos formativos \neq modelos científicos de medición

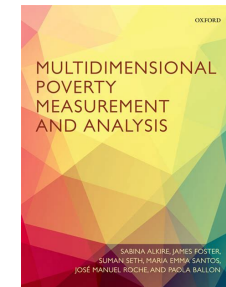
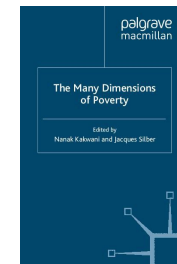
El reloj (instrumento) \neq el tiempo mismo (fenómeno)

¿Cambian los indicadores, cambia la definición?

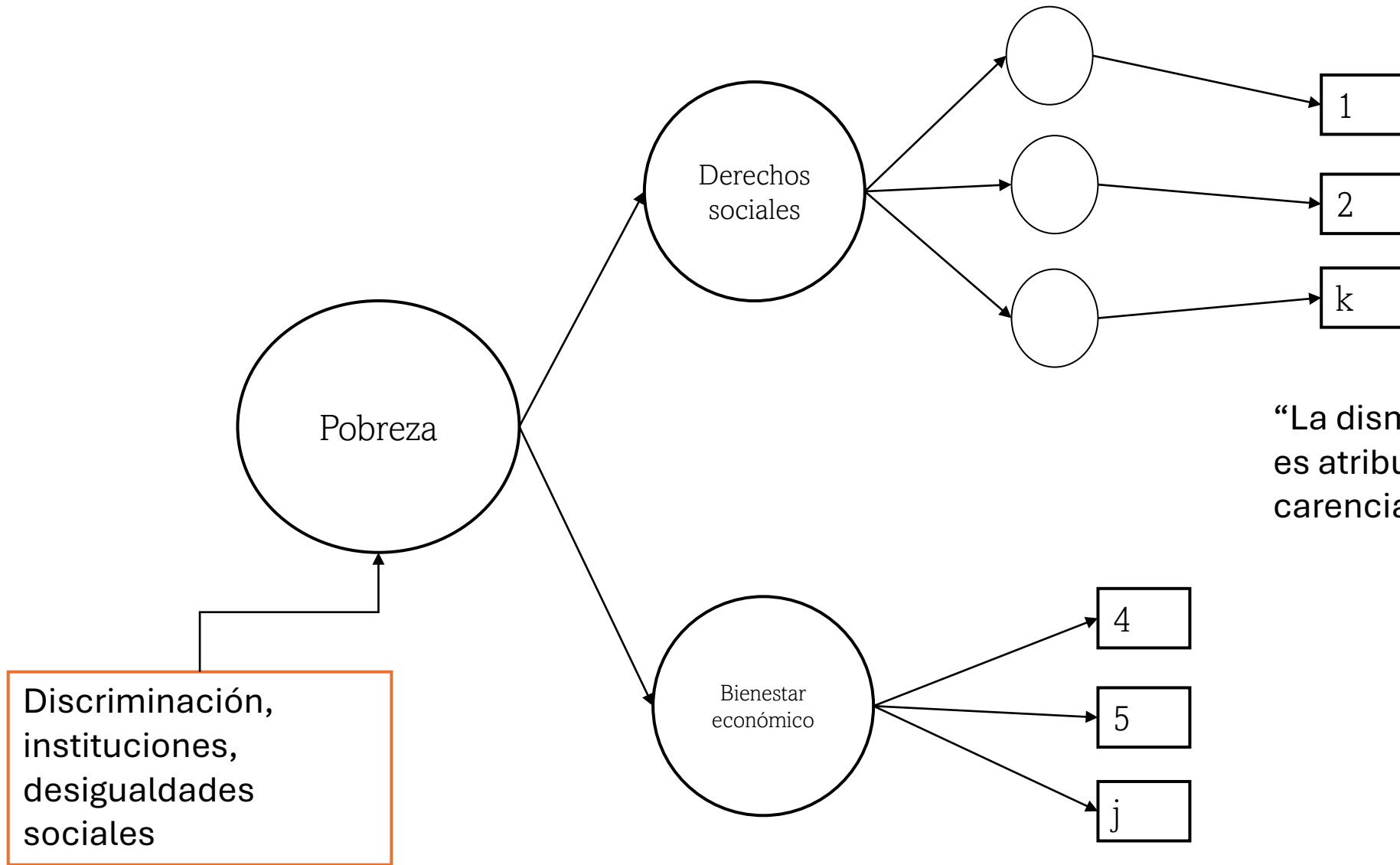
¿Cambian los indicadores, cambian las causas?

¿LDGS? ¿NBI?

La pobreza son los indicadores/dimensiones que la componen

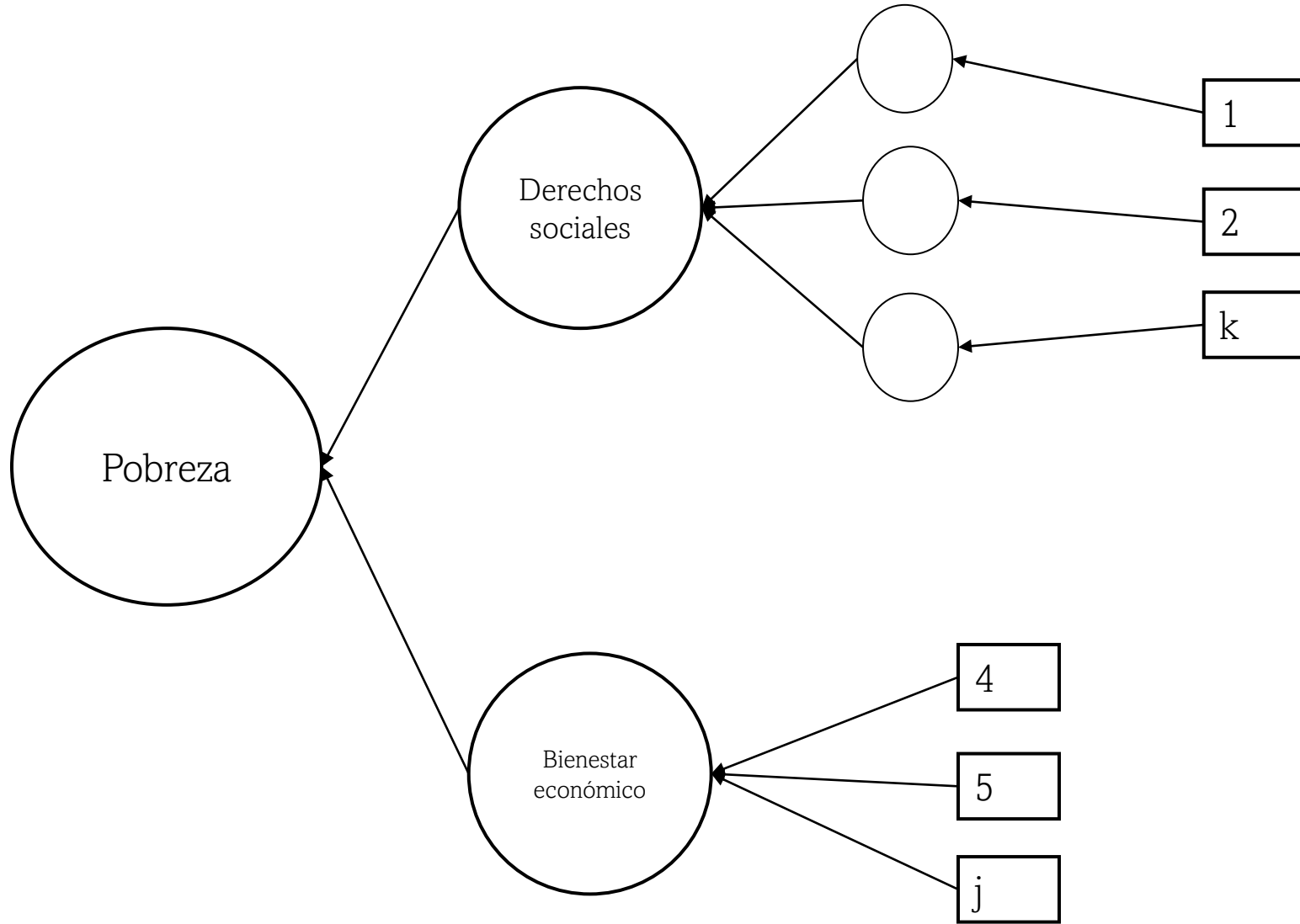


Modelo mexicano



“La disminución de la pobreza es atribuible a que las carencias mejoraron”

Modelo mexicano



“La disminución de la pobreza es atribuible a que las carencias mejoraron”

¿Causas = consecuencias?

Validez de “cara” (face validity)

- La interpretabilidad de los scores según la población en estudio
- Es decir, la población nos indica si lo que estamos midiendo les hace sentido desde sus vivencias
- Se usa para producir cuestionarios y analizar sus resultados
- Últimamente ha ganado eco por el fracaso que han tenido medidas exclusivamente diseñadas desde el escritorio



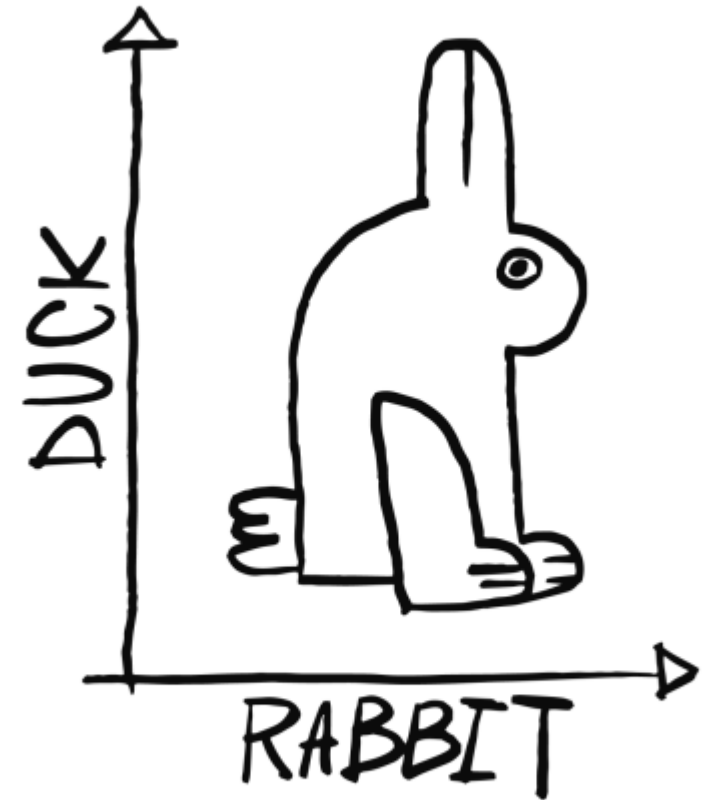
Métodos: Validez de cara

- La validez de cara requiere métodos mixtos:
Cualitativos y Cuantitativos
- Cualitativos: Entrevistas a profundidad,
grupos focales e incluso etnografía
- Cuantitativos

Condición necesaria: validez de cara

Se llega a disputar su carácter de condición necesaria:

- Preferencias adaptativas



Ejemplos

¿Qué bienes, servicios y actividades son esenciales para la vida?

¿Qué componentes deberían ser parte de la canasta básica?

¿Qué manifestaciones tiene una economía en recesión?

¿Hacen sentido para los afectados por la pobreza, la recesión o la inflación?

% thinking items a necessity (Sex and Age of Respondent)	Sex		Age group	
	Male	Female	<24 yrs	65+
A visit to a health facility when ill and all the medication prescribed to treat the illness	97	98	98	97
Three meals a day	96	95	95	96
Two sets of clothing	94	93	93	92
Toiletries to be able to wash every day (e.g. soap, hairbrush/comb)	93	93	94	90
All fees, uniform of correct size and equipment required for school e.g. books, school bag, lunch/lunch money, etc.	89	88	88	86
Own blanket	85	85	87	83
Own bed	81	81	82	79
Two pairs of properly fitting shoes, including a pair of all-weather shoes	80	79	82	74
Own room for children over 10 of different sexes	78	75	74	77
Books at home suitable for their age (including reference and story books)	72	71	72	69
Some new clothes (not second hand or handed on/down)	70	69	73	64
Bus/taxi fare or other transport (e.g. bicycle) to get to school	69	68	69	66
To be able to participate in school trips or events that cost money	68	69	70	66
A desk and chair for homework for school aged children	57	54	56	54
Presents for children once a year on special occasions, e.g. birthdays, Christmas, Eid	55	53	57	52
Educational toys and games	54	52	57	50
Some fashionable clothes for secondary school aged children	38	37	43	35
Own cell phone for secondary school aged children	23	21	26	24

Ejemplo: Medición oficial mexicana

Elección del umbral oficial

Educación: 10%

Paredes: 1 %

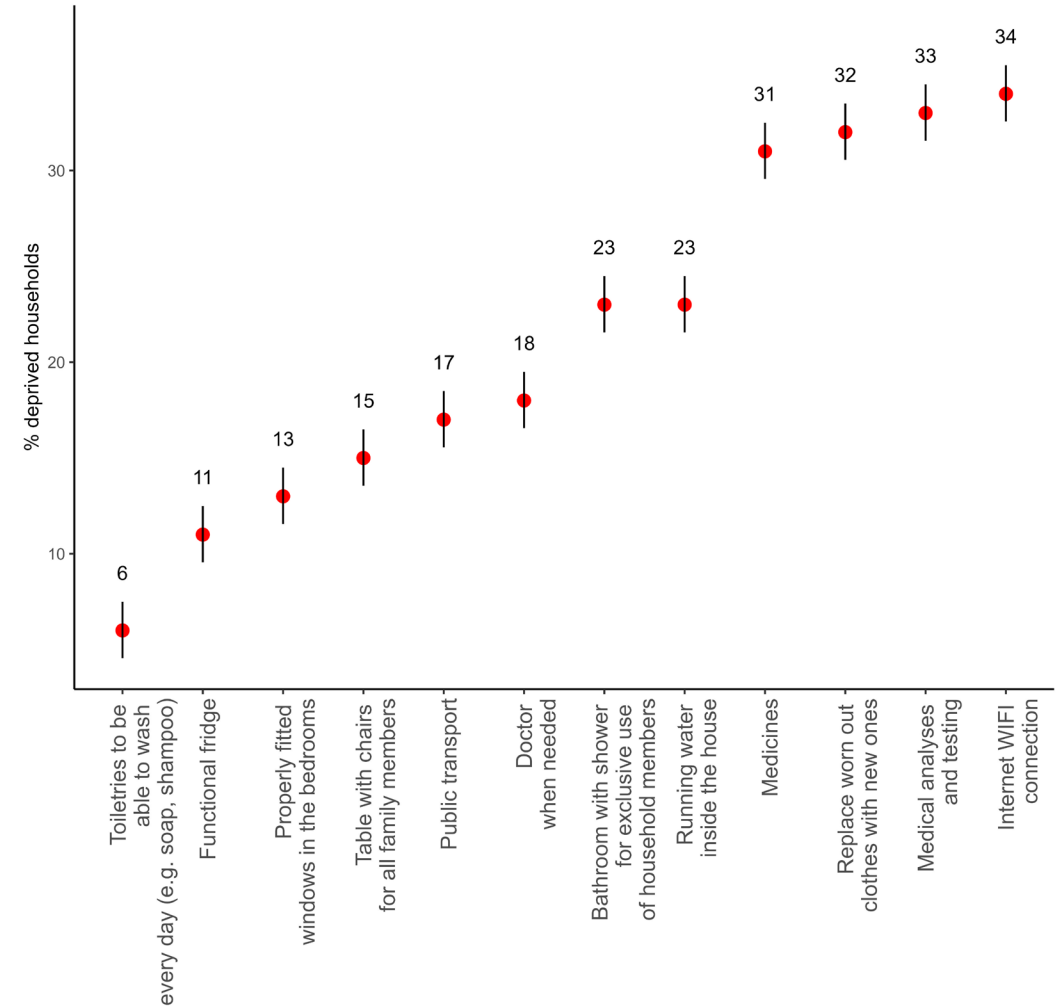
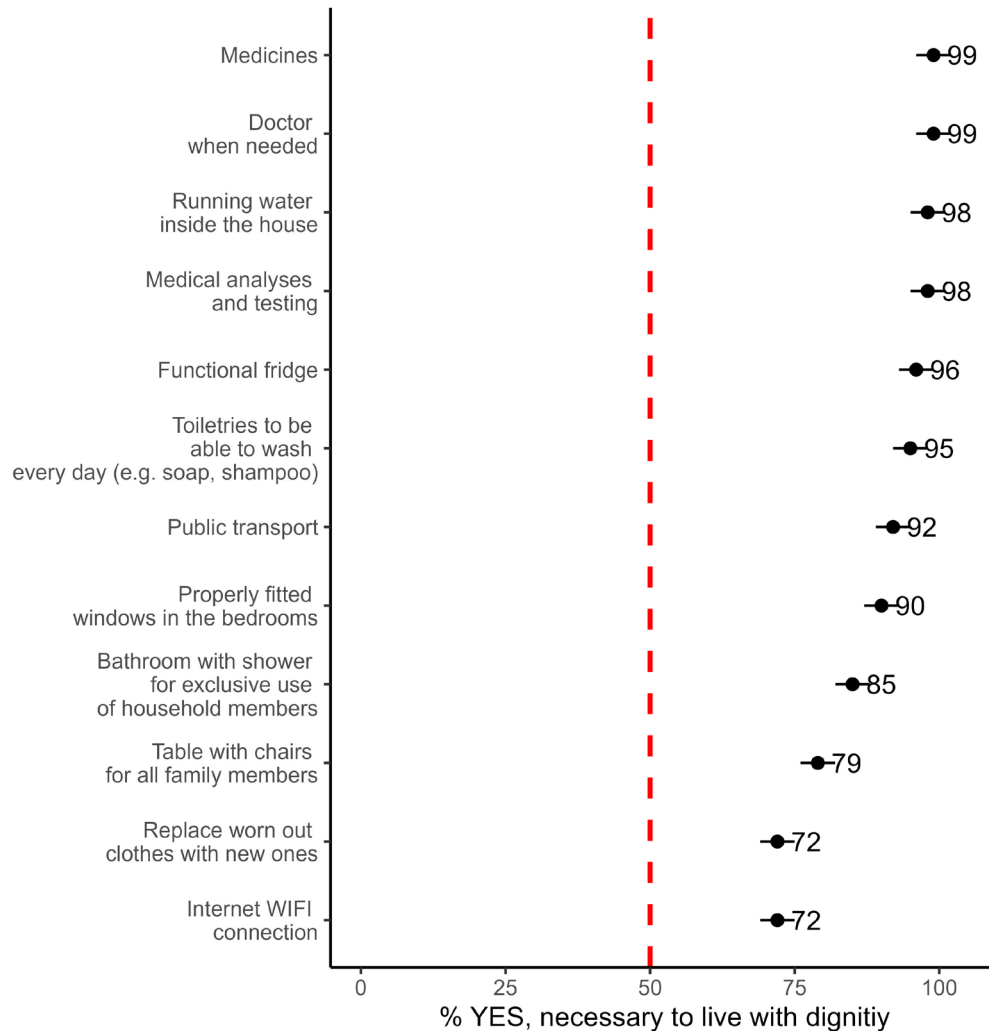
Techos: 4%

Agua: 9%

Social deprivation indicators	CONEVAL's estimates (%)	Consensual approach
1. Educational gap	19	21
2. Access to health care services	22	22
3. Access to social security	61	65
4. Quality and living space of the dwelling	14	35
5. Access to basic services in the dwelling	21	52
6. Access to food	23	23
7. Access to information	n/a	68

EDUMP 2017

Ejemplo: Privación relativa

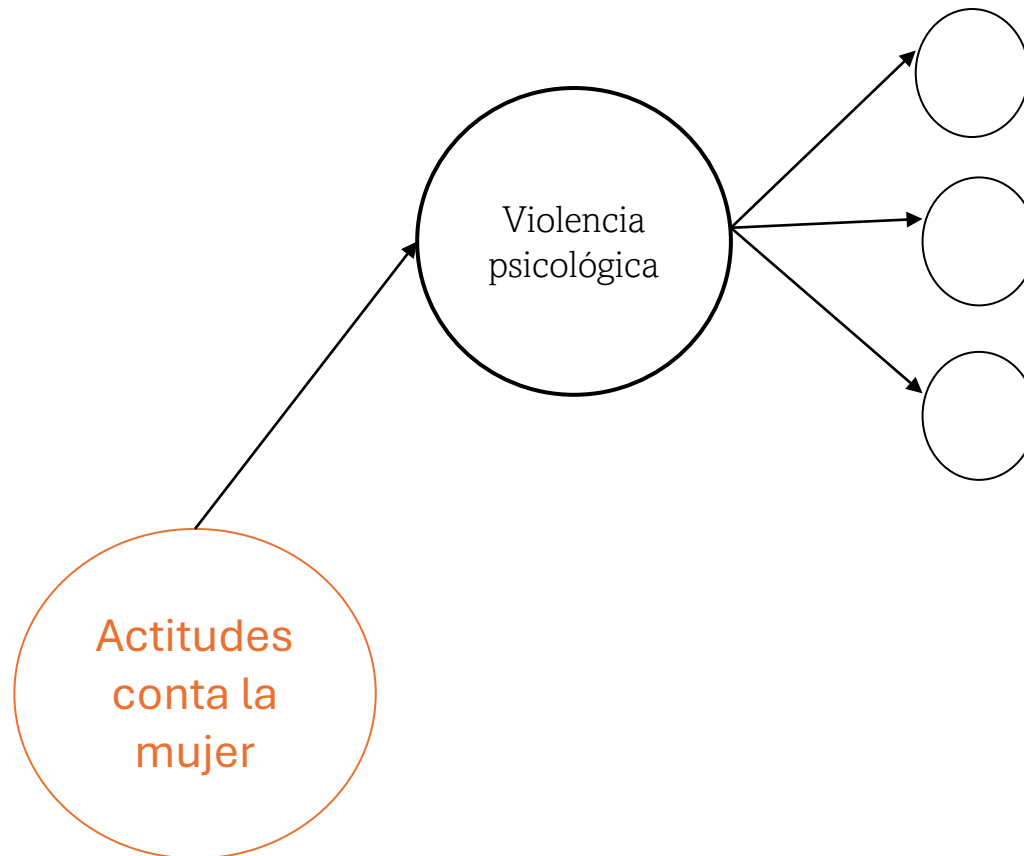


Validez de criterio: Sus inicios

- En los 30s: Los conceptos son (a) convenientes, creados, constructos explicativos y (b) son inobservables pero inferidos a través de manifestaciones
- La primera forma de validación de los scores consistía en su capacidad de predecir algún “criterio”

Por ejemplo:

Considere el siguiente modelo

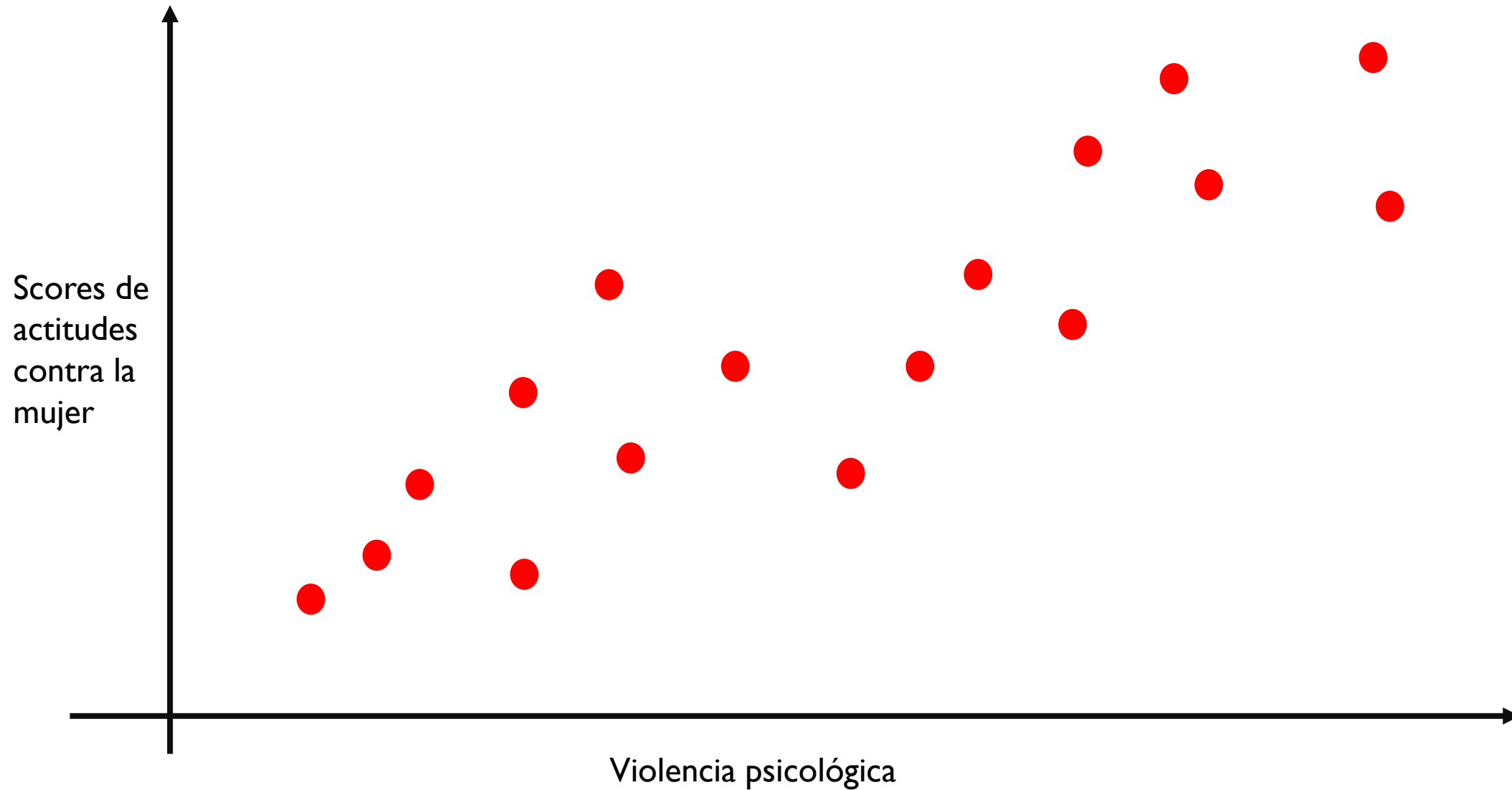


Teoría:

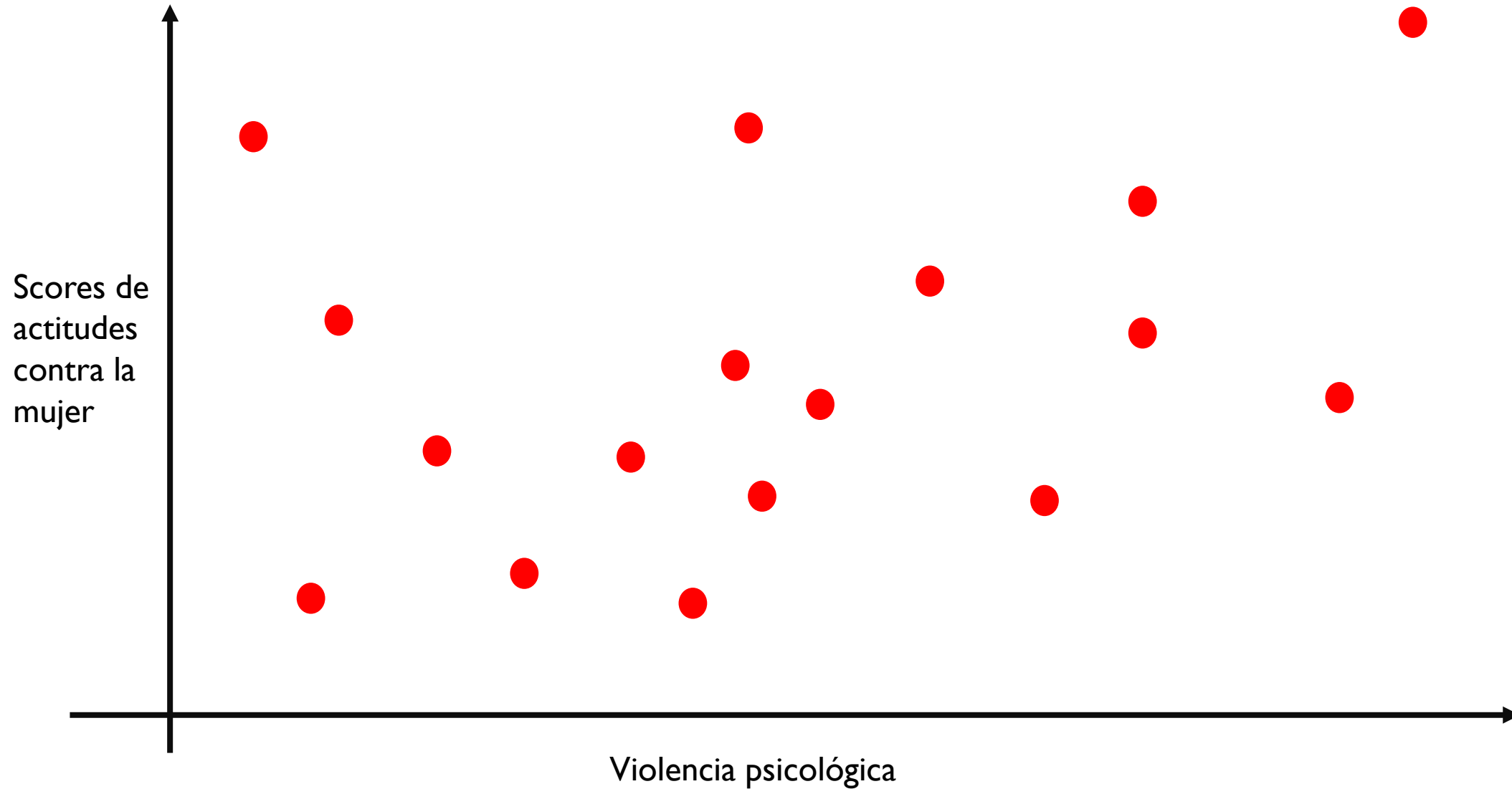
Las normas sociales justifican ciertos juicios sobre lo que es permisible en las relaciones de género. Estas normas se traducen en **actitudes contra la mujer**.

Cuando ciertas actitudes negativas se asientan, el riesgo de perpetrar **violencia psicológica** aumenta.

Validez de criterio



Validez de criterio



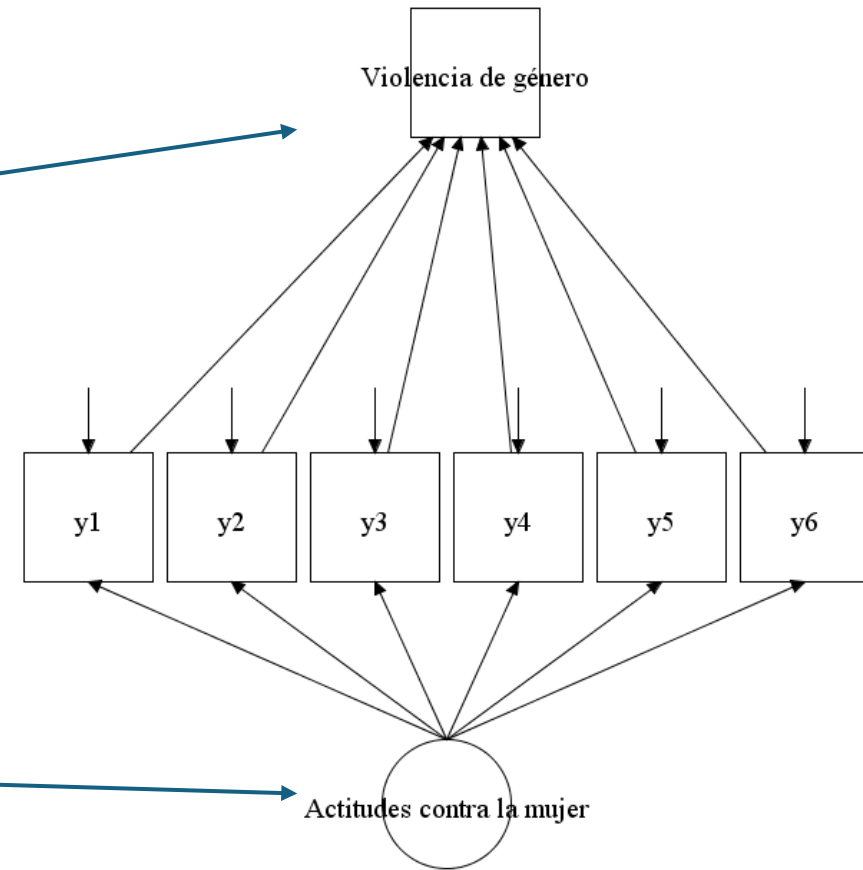
Validación de criterio a la antigua

Lo que se hacía era correlacionar los indicadores con alguna variable auxiliar

Esto llevo a pensar en “proxies”

Uno tenía correlaciones / pendientes para cada indicador

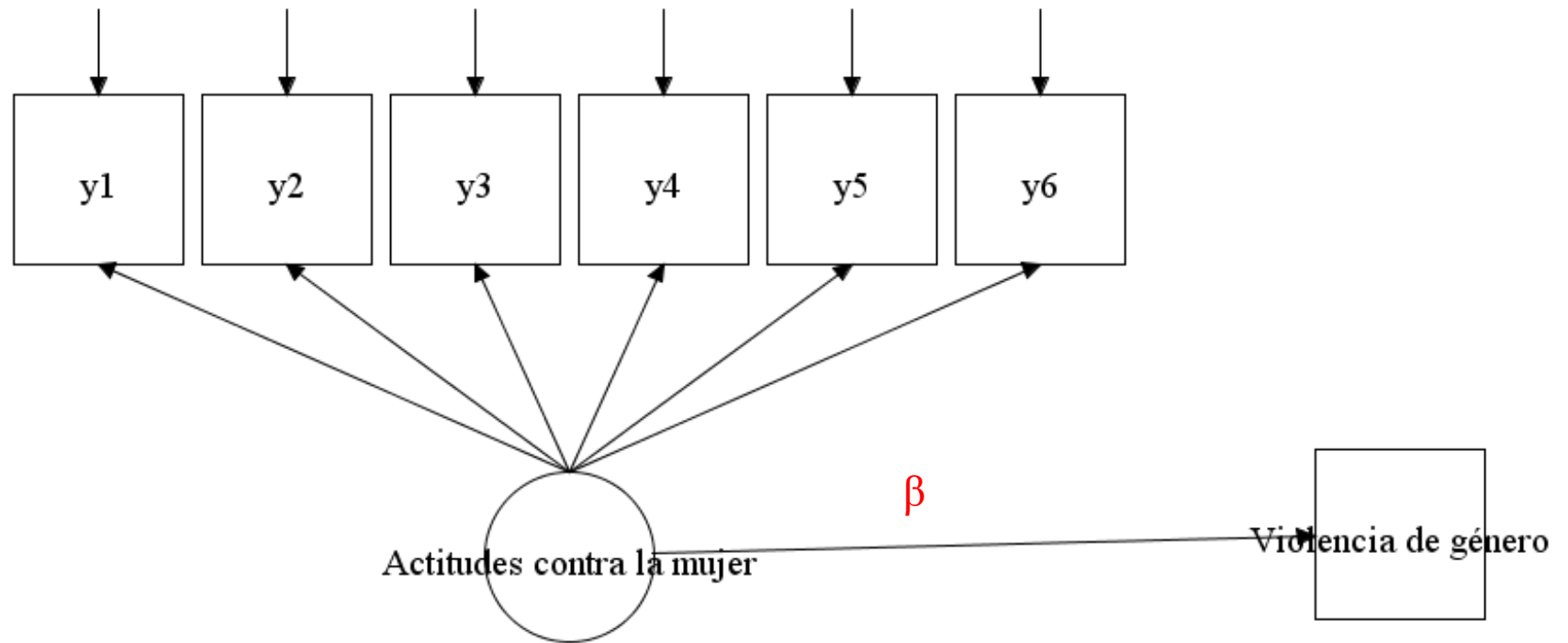
Antes era muy difícil calcular el valor de los scores latentes



SEM: Validez de criterio o predictiva

Con SEM podemos
usar directamente
los scores latentes

Un coeficiente nos
lo dice todo

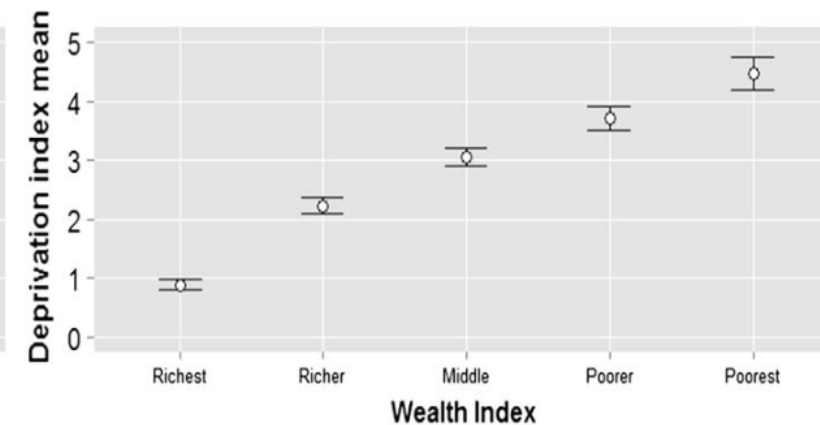
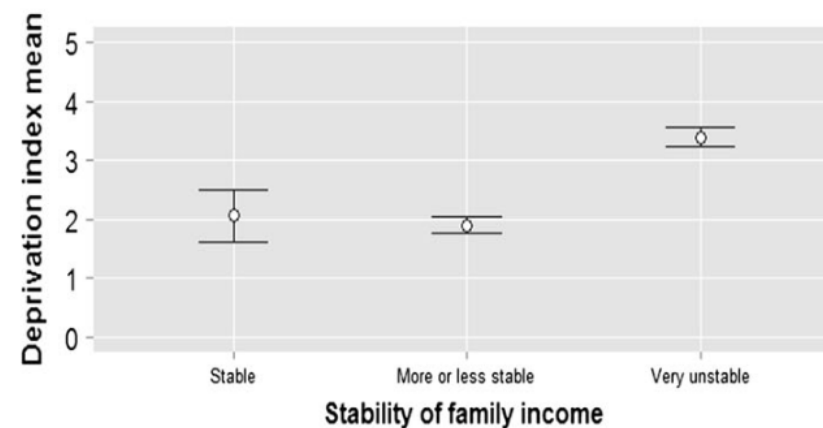
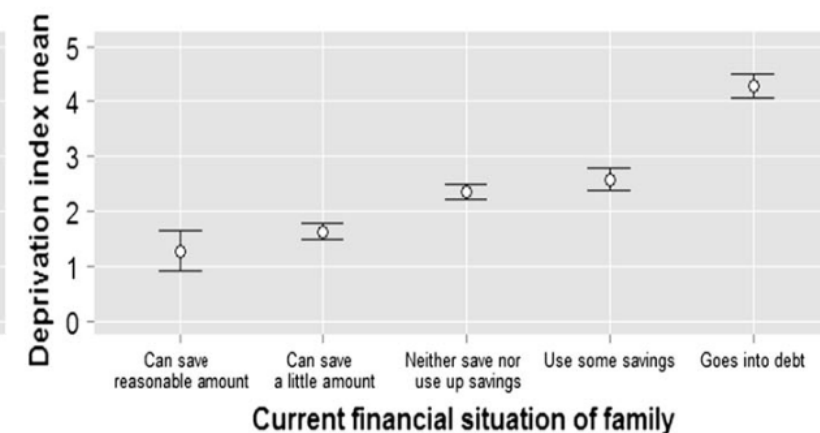
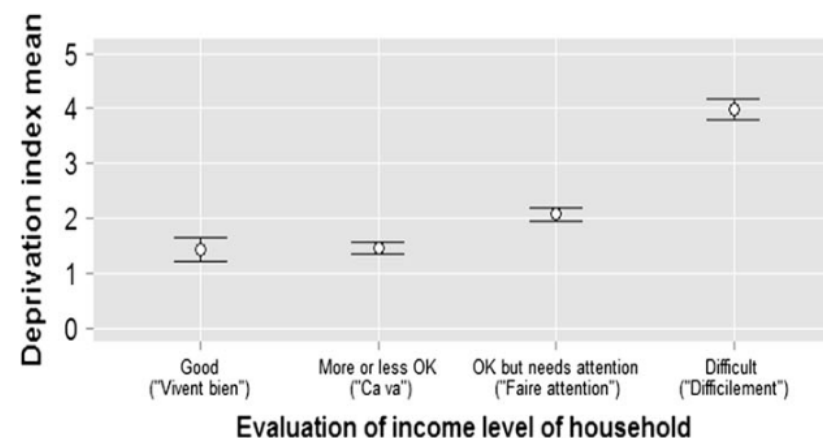


Validez de criterio: Descriptivos

S. Nandy, M. Pomati

Siempre hay que trabajar con hipótesis. Esto no es un análisis de sensibilidad.

¿Cuáles son los cambios en gradientes esperados?



Validez criterio

1. Head of household's education level (scored from 1 to 8 – from 'degree' to 'no formal education') – **HHEd**
2. Expenditure poverty using the official measure (1 poor, 0 not poor) – **Poor**
3. Head of household's economic activity ('working in the cash economy' vs 'subsistence/family worker or unemployed') – **HHact**
4. Subjective poverty (scored from 1 to 5 - from 'very rich' to 'very poor') – **Sub_pov**
5. International Socio-Economic Index (ISEI) of occupational status– a widely-used measure of occupational status. Lower scores indicate higher occupational prestige – **ISEI**.

TABLE A1.8: LOGISTIC REGRESSION VALIDITY TESTS FOR CHILDREN AND HOUSEHOLD DEPRIVATION ITEMS

Children and household items + 2 adult items concerning children		HHEd	Poor	HHact	Sub_pov	ISEI
1	Child: Three meals a day	1.5	5.5	1.8	3.5	28.7
2	Child: Two pairs of properly-fitting shoes	1.6	6.9	2.3	2.9	44.6
3	Child: Toiletries to be able to wash every day	1.4	3.3	1.5	2.7	15.2
4	Child: Books at home for their age	1.3	2.1	1.3	1.6	7.1
5	Child: Some new clothes	1.4	4.1	1.8	2.6	18.0
6	Child: Educational toys and games	1.2	2.4	1.3	1.4	4.1
7	Child: A visit to health facility when ill and all prescribed medication	1.2	1.7	1.1	2.0	6.3
8	Child: Own bed	1.5	5.1	1.8	2.5	16.8
9	Child: Own blanket	1.5	5.1	1.8	2.5	19.1
10	Child: Two sets of clothing	1.4	2.9	1.6	2.6	17.5
11	Child: Presents for children once a year on special occasions	1.3	3.1	1.7	2.1	8.9
12	Child: All school fees, uniforms of correct size and equipment	1.3	2.1	1.4	1.7	7.8
13	Child: To be able to participate in school trips	1.2	1.8	1.2	1.5	5.1
14	Child: A desk and chair for homework	1.2	1.6	1.1	1.3	3.0
15	Child: Bus/taxi fare or other transport	1.2	1.7	1.3	1.5	3.9
16	Child: Own room for children over 10 of different sexes	1.1	1.3	1.0	1.3	1.2
17	Household: Enough money to repair or replace worn-out furniture	1.4	4.1	1.9	2.9	19.0
18	Household: Enough money to repair or replace electronic goods	1.2	2.0	1.3	1.6	5.7
19	Household: To be able to make savings for emergencies	1.3	3.4	1.8	2.6	10.5
20	Household: Enough money to repair a leaking roof for main living q.	1.3	2.4	1.5	2.0	9.7
21	Household: To have own means of transport	1.2	2.1	1.7	1.9	3.2
22	Household: Enough money to replace broken pots and pans	1.3	2.9	1.8	2.3	11.6
23	Adult: Enough money to pay school fees	1.3	1.9	1.3	1.7	5.6
24	Adult: Enough money to take the child to a medical facility when sick	1.3	3.0	1.6	2.4	10.7

Esto ayuda
mucho, pero
¿cuáles
serían los
problemas
potenciales?

Appendix 3: Item validation

	Difficult compared to good or more or less OK (relative risk with 95 % CIs)
Validator 1—evaluation of household income status	
Number of meals every day	14.7 (10.8–20.1)
Consumption of cereals and tubers every day	9.4 (6.9–12.8)
Clothing	6.3 (5.0–7.9)
Shoes	6.0 (4.8–7.5)
A good meal on festivities/celebrations (Sunday, ceremony, etc.)	5.7 (4.3–7.5)
Consumption of meat or fish every day	5.6 (4.5–6.9)
Cleanliness/personal hygiene	4.1 (3.4–5.1)
Housing	3.8 (3.3–4.4)
Care in case of sickness	3.6 (3.1–4.3)
Self-care products in the house	3.5 (3.0–4.1)
Education for children	2.8 (2.3–3.2)
Availability of transport	2.5 (2.3–2.7)
Availability of leisure	2.2 (2.0–2.5)
Furniture in the house	2.2 (2.0–2.4)
Availability of drinking water	1.9 (1.7–2.1)
Availability of electricity	1.6 (1.5–1.7)

¿Qué hay de esto?

Ingreso per capita	Coef.	P>t	[95% Conf.	Interval]
Pisos	-131	0.361	-412	150
Techos	43	0.092	-7	92
Paredes	202	0.238	-133	538
Hacinamiento	-735	0.000	-867	-603
Tenencia	-76	0.352	-236	84
Agua	-179	0.009	-314	-44
Saneamiento	-362	0.000	-493	-232
Combustible	-971	0.000	-1106	-837
Bienes durables	-143	0.065	-296	9
Educacion basica inf	-244	0.056	-493	6
Rezago educativo	-78	0.557	-338	182
Educacion adultos	-422	0.000	-610	-234
Desempleo	-393	0.000	-513	-272
Proteccion social	-1445	0.000	-1704	-1186
_cons	4714.443	0.000	4459.285	4969.601

Es manifestación
o causa de
pobreza.

Parecen tener la misma fuente de variación pero porque son parte de la misma!

No es un índice de pobreza sino índice de causas y consecuencias de la pobreza

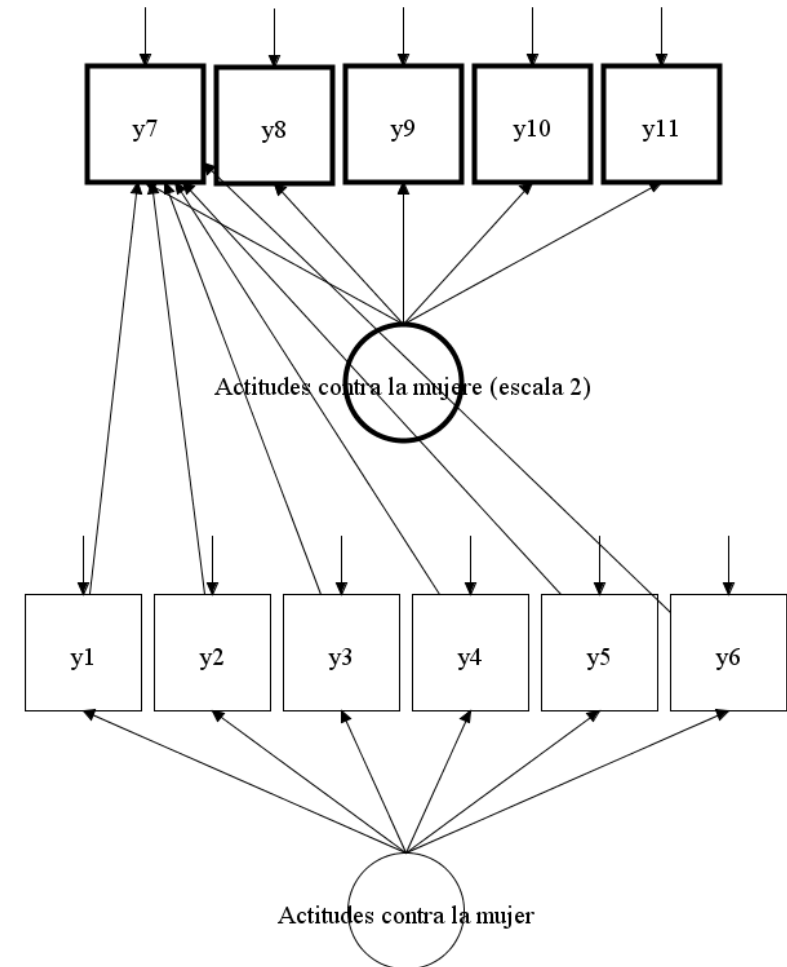
Validez concurrente

- Spearman 1904: En aquella época la distinción entre confiabilidad y validez no era tan clara
- Spearman tenía distintas medidas de inteligencia y habilidad musical
- La idea es similar pero el foco es son los scores globales entre distintos tests

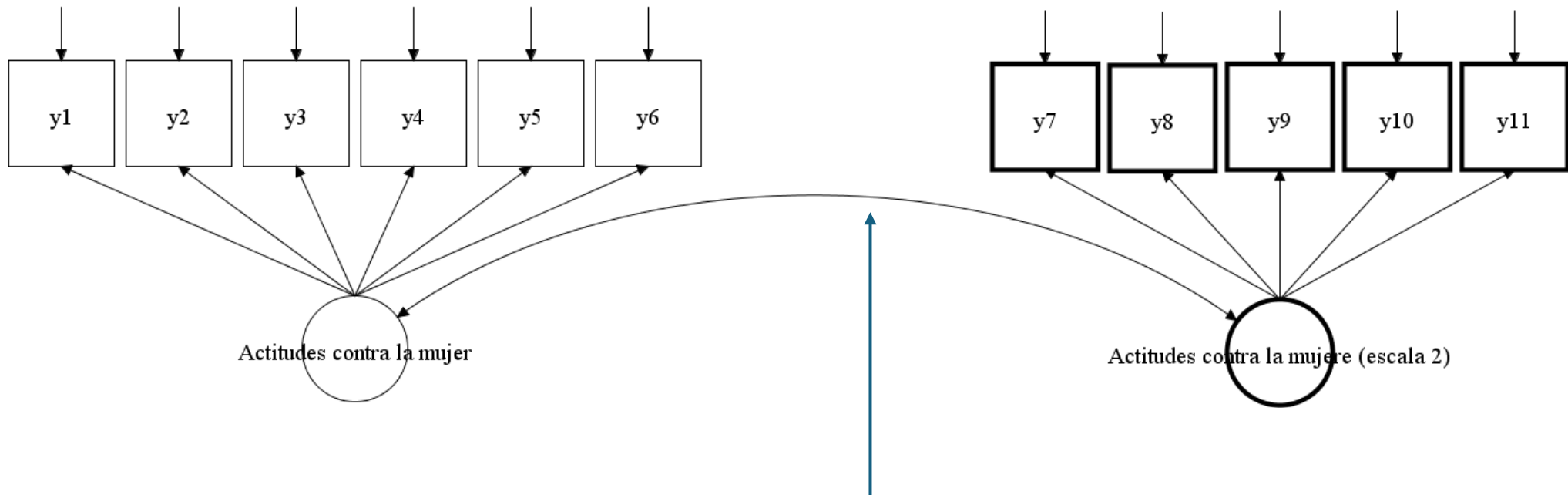
Validación concurrente a la antigua

Antes se tomaban las correlaciones entre distintos componentes de diferentes tests

Si se dan cuenta es similar a confiabilidad porque se piensa que los componentes miden el mismo fenómeno

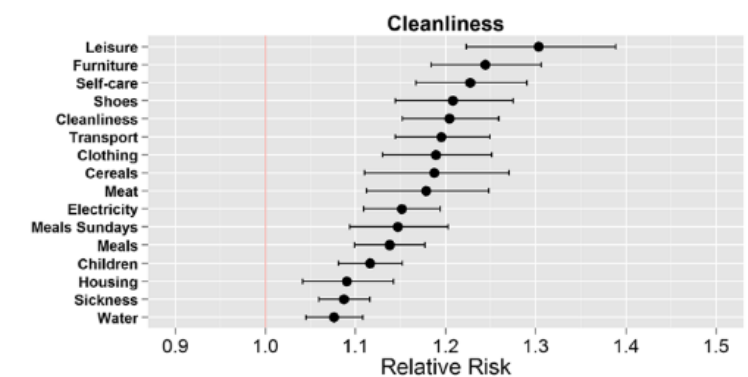
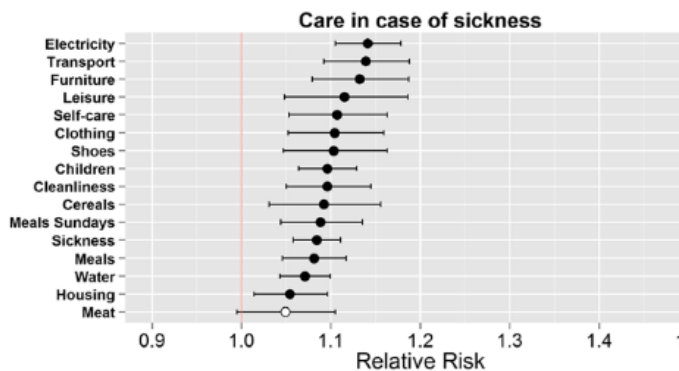
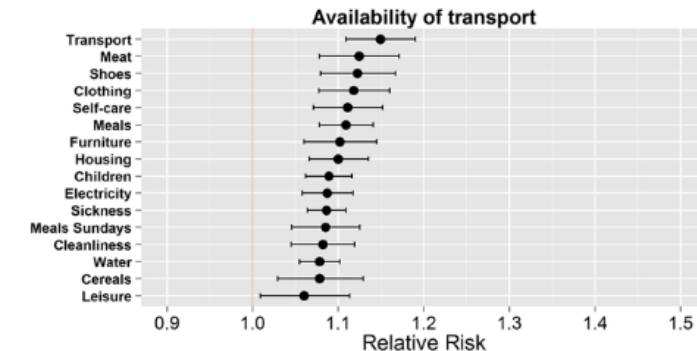
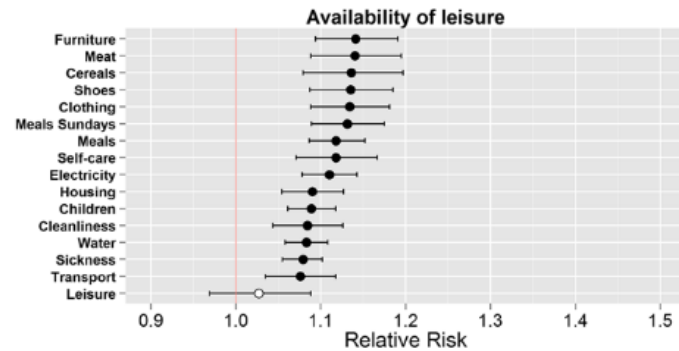
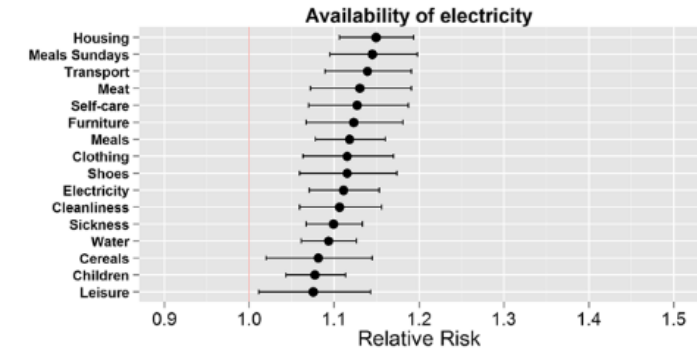
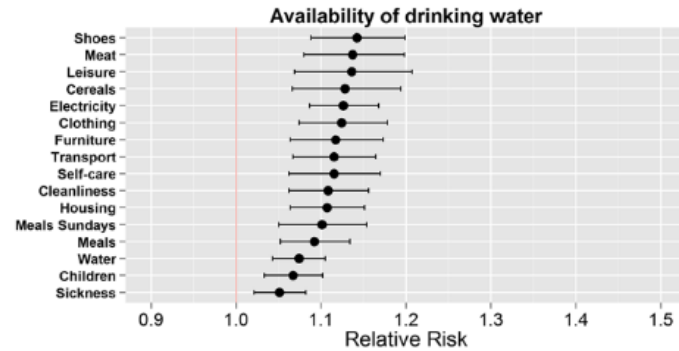


SEM: Validación concurrente



Con un coeficiente podemos examinar la validación concurrente. PERO... Son latentes con latentes.

Ejemplo: Validez concurrente

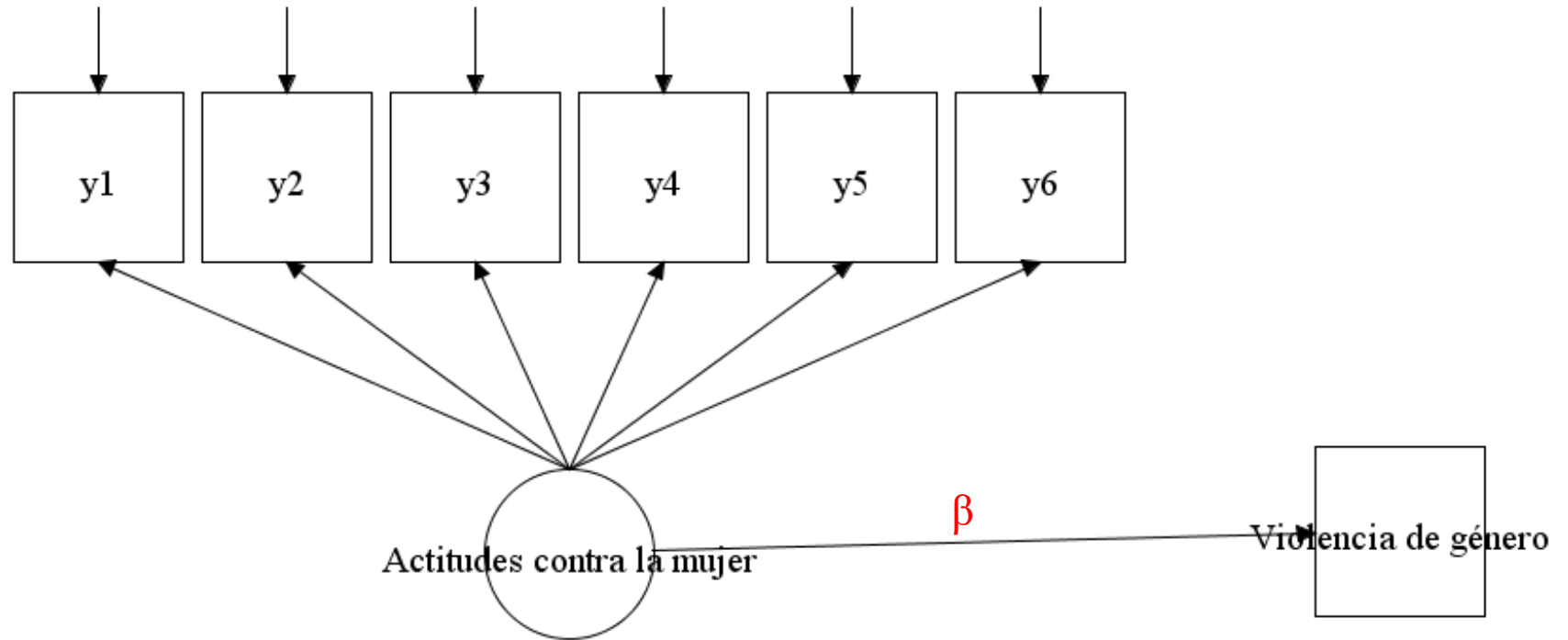


Validez de constructo

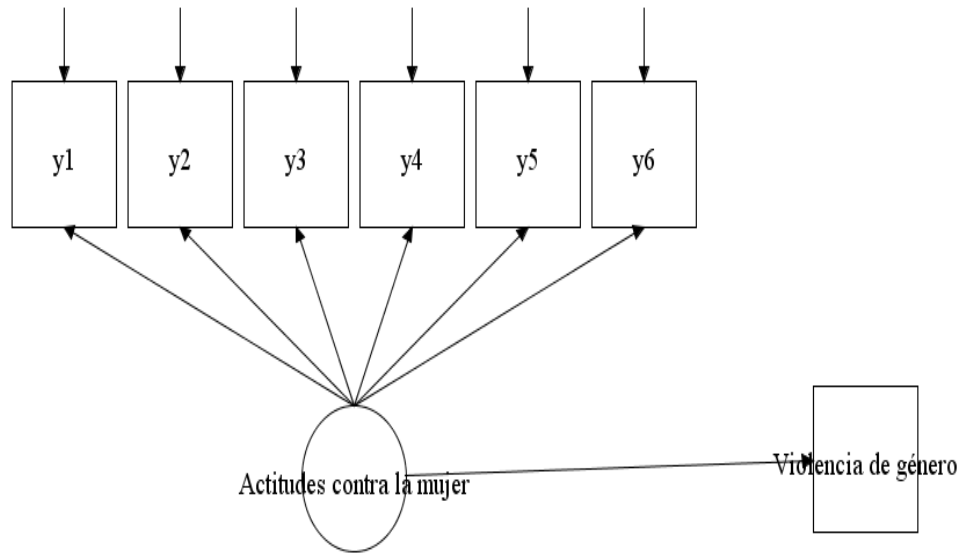
- Durante décadas hubo varios tipos de validez
- En 1954 la sociedad americana de psicología enlistó los tipos de validez y añadió un cuarto: constructo
- En una serie de contribuciones Cronbach, Meehl y Messick anidaron todos los tipos de validez en la validez de constructo
- ¿Por qué?
- Al final todos los tipos de validez ven desde distintos ángulos si los scores pueden interpretarse como queremos

Validez de constructo

No sólo se
preocupa por β
sino por la
estructura de la
medición

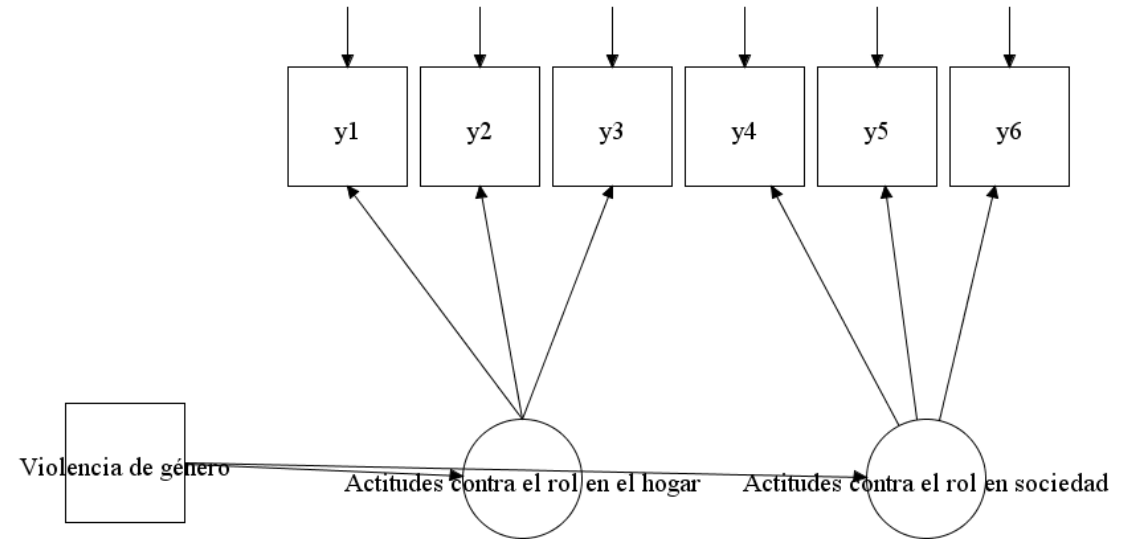


Validez de constructo



TLI = .8
RMSEA= .07
CFI = .76

VS



TLI= .94
RMSEA= .04
CFI= .94

Validez de constructo

Figure 2 UNDP Multidimensional Poverty Index Hierarchical Model and Weightings

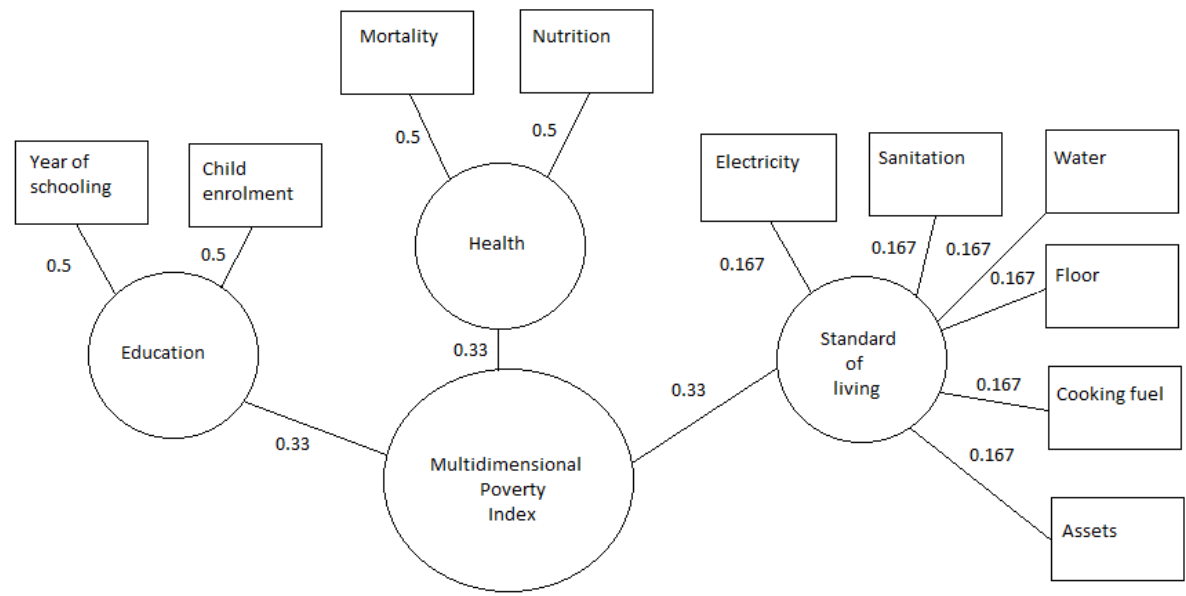
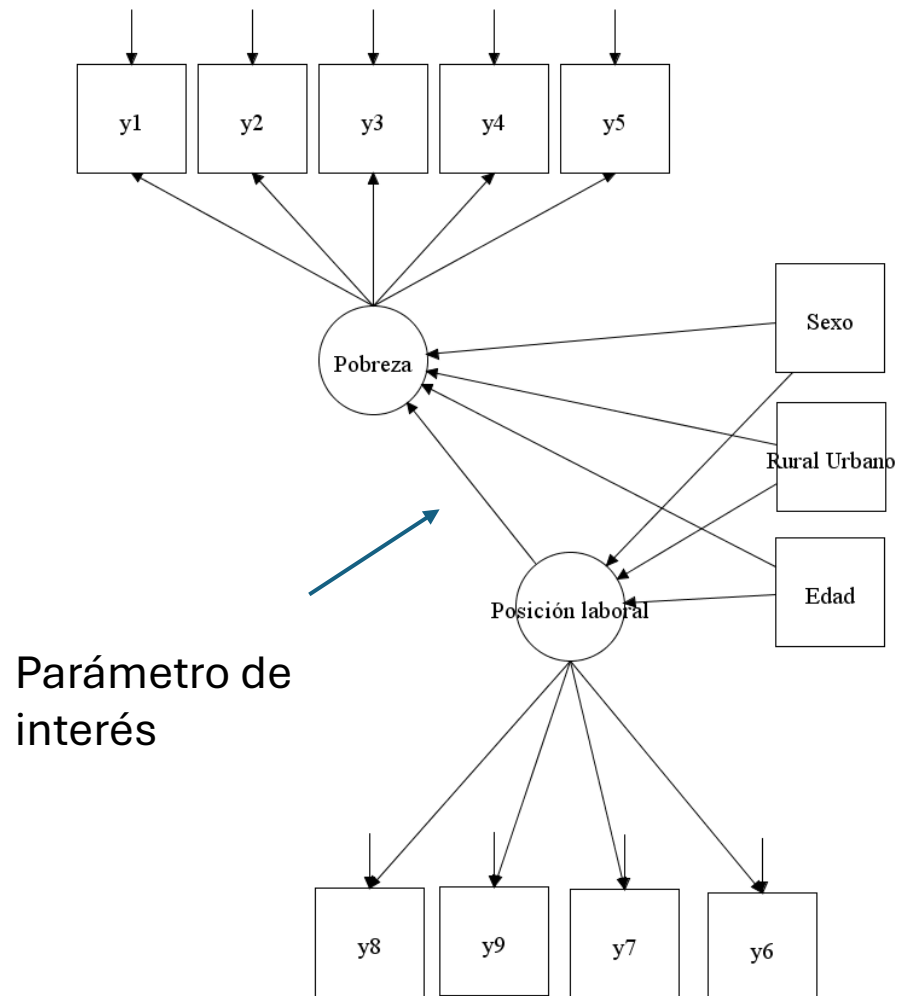


Table 3. Confirmatory Factor Analysis. Model Fit statistics

		H1	H2	H3	H4
Country	Statistic	Fixed item loadings	Fixed dimension loadings and free item loadings	Full free weights	Unidimensional
Uganda	TLI	0.73	0.94	0.85	0.94
	RMSEA	0.08	0.04	0.06	0.04
	TLI	0.78	0.83	0.68	0.97
Benin	RMSEA	0.06	0.06	0.08	0.02
	TLI	0.79	0.91	0.92	0.94
Cameroon	RMSEA	0.08	0.05	0.05	0.04
	TLI	0.67	0.88	0.75	0.92
Congo	RMSEA	0.06	0.03	0.05	0.03
	TLI	0.39	0.96	0.88	0.97
Congo DR	RMSEA	0.10	0.03	0.04	0.02
	TLI	0.63	0.82	0.72	0.88
Ghana	RMSEA	0.06	0.04	0.05	0.03
	TLI	0.92	0.89	0.85	0.97
India	RMSEA	0.04	0.04	0.05	0.02
	TLI	0.89	0.94	0.92	0.97
Pakistan	RMSEA	0.04	0.03	0.04	0.02
	TLI	0.50	0.79	0.79	0.91
Nigeria	RMSEA	0.06	0.04	0.04	0.03
	TLI	0.73	0.83	0.69	0.94
Guinea	RMSEA	0.07	0.05	0.07	0.03

SEM y Validez

- La flexibilidad de SEM permite hacer en un paso modelos causales:



Si tengo un buen modelo teórico (*distinción entre causas y consecuencias*), puedo hacer ejercicios de validación más completos y unificados

MODEL:

Pobreza BY y_1 y_2 y_3 y_4 y_5 ;

Posición laboral BY y_6 y_7 y_8 y_9 ;

Pobreza ON Posición laboral Sexo Edad Rural Urbano;

Posición laboral ON Sexo Edad Rural Urbano;

Validez y complementariedad: EMSA y ELCSA

Why identifying households by degree of food insecurity matters for policymaking

Rafael Pérez-Escamilla ^a  , Mireya Vilar-Compte ^b , Pablo Gaitan-Rossi ^b 

Show more 

+ Add to Mendeley  Share  Cite

<https://doi.org/10.1016/j.jgfs.2020.100459>

[Get rights and content](#)

A Model-Based Approach to Identify Classes and Respective Cutoffs of the Brazilian Household Food Insecurity Measurement Scale

Michael E Reichenheim, Gabriela S Interlenghi, Claudia L Moraes, Ana M Segall-Corrêa, Rafael Pérez-Escamilla, Rosana Salles-Costa 


The Journal of Nutrition, Volume 146, Issue 7, July 2016, Pages 1356–1364, <https://doi.org/10.3945/jn.116.231845>

Published: 08 June 2016 **Article history** 



PDF

 Split View

 Cite

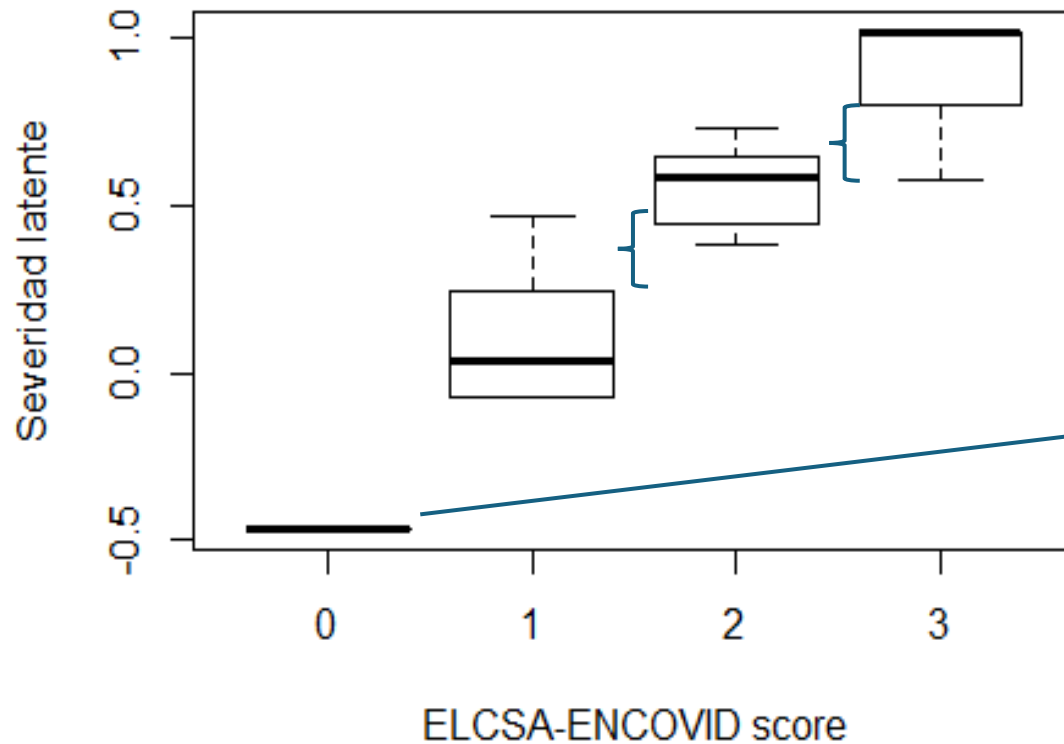
 Permissions

 Share 

Conclusion: Although roughly classifying EBIA as in previous studies, the current approach suggests that, in terms of raw score, households endorsing only one item of the scale would be better classified by being placed in the same stratum as those remaining negative on all items.

¿Es grave que no se sostenga el Rasch?

Severidad latente por grupo. IRT Dos parámetros. Ad



No hay buena distinción en severidad de los grupos de inseguridad 1 (leve) y 2 (moderada)

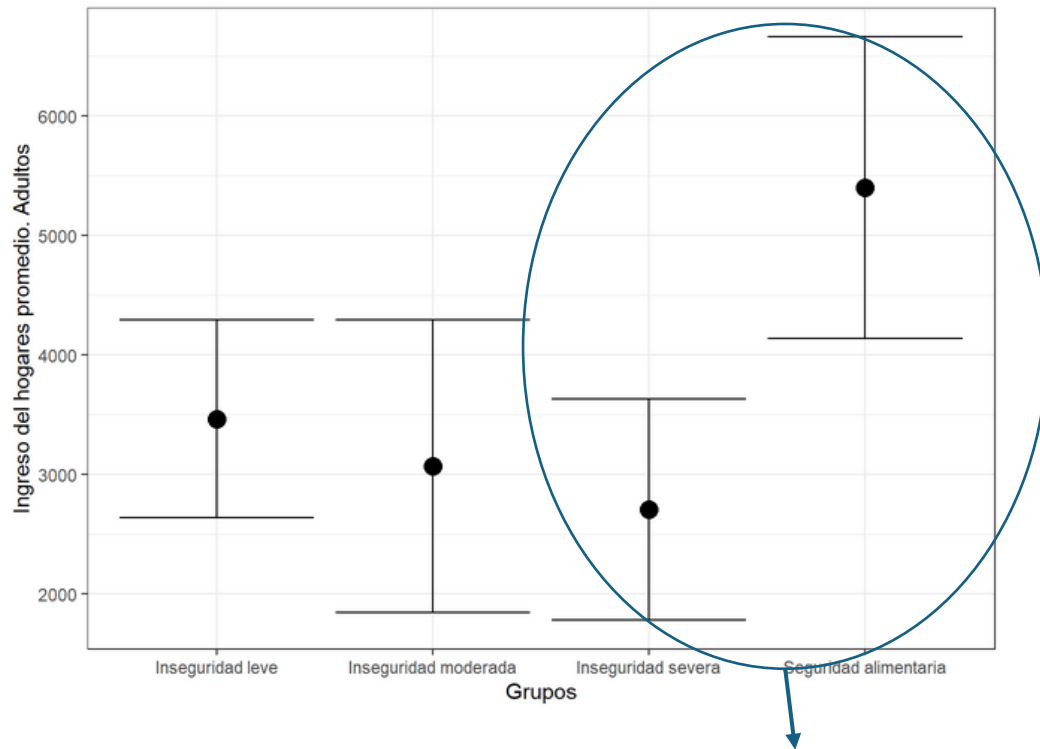
El grupo de claramente seguros parece ser el primer grupo.

No parece que el grupo 1 sea próximo al 0 y, por tanto, que puedan agruparse!

Esto es grave si se hace la construcción manual de grupos y puntajes!

Validez de criterio de los cuatro grupos

Validez de criterio de los 4 grupos



Las conclusiones sobre cambios en inseguridad alimentaria se basan en agrupar estos dos conjuntos.

El resultado es que nuevamente no parece haber evidencia que sustente a los 4 grupos.

En el artículo que acompaña la ENCOVID-19 hay conclusiones un tanto distintas

Sin embargo, esto ocurre con el ingresos de la ENCOVID-19 pero también (aunque de manera menos aguda) con la EMSA y la ENIGH.

Nunca confundan validez con confiabilidad

Validez estadística de la Escala Mexicana de Seguridad Alimentaria y la Escala Latinoamericana y Caribeña de Seguridad Alimentaria

Paloma Villagómez-Ornelas, MPD,⁽¹⁾ Pedro Hernández-López, L en Econ,⁽¹⁾ Brenda Carrasco-Enríquez, Act,⁽²⁾
Karina Barrios-Sánchez, L en C de la Com,⁽¹⁾ Rafael Pérez-Escamilla, D en Nut,⁽³⁾ Hugo Melgar-Quirón, D en C Méd.⁽⁴⁾

Scale (ELCSA). **Materials and methods.** Validity tests were conducted in order to verify that both scales were consistent instruments, conformed by independent, properly calibrated and adequately sorted items, arranged in a continuum of severity. The following tests were developed: sorting of items; Cronbach's alpha analysis; parallelism of prevalence curves; Rasch models; sensitivity analysis through mean differences' hypothesis test. **Results.** The tests showed that both scales

¿Cuál es la
situación
actual en
México en
medición?

Hacia una encuesta social general

Resultados de los trabajos del
Grupo Técnico Especializado
en la Medición del Ingreso y
el Bienestar

Raymundo M. Campos Vázquez,
Gerardo Leyva Parra
y Graciela Teruel Belismelis

EL COLEGIO DE MÉXICO
UNIVERSIDAD IBEROAMERICANA
INSTITUTO NACIONAL DE ESTADÍSTICA
Y GEOGRAFÍA

28 Truncamiento, subreporte, no respuesta y otras fuentes de error en la medición del ingreso

Mecanismos para minimizar el error total de las encuestas y mejorar la estimación de la distribución del ingreso

Raymundo Campos,¹ Gerardo Esquivel y Delfino Vargas

¿Cuál es la verdadera distribución del ingreso? En México, el método por excelencia para estimar la distribución del ingreso ha sido encuestar hogares. Sin embargo, existen diferentes errores estadísticos que impiden afirmar que el ingreso reportado en las encuestas de los hogares es el verdadero. Tienen que ver con la cobertura, el muestreo, la tasa de respuesta y la precisión de la información, y tienen implicaciones importantes para la investigación y la política pública. Aquí se proponen diferentes mecanismos para minimizar el error total de las encuestas y mejorar el cálculo del ingreso. Estos abarcan la inferencia de datos, el reemplazo, la imputación, la reponderación con datos de la misma encuesta y el uso de datos administrativos como solución a los errores de medición.

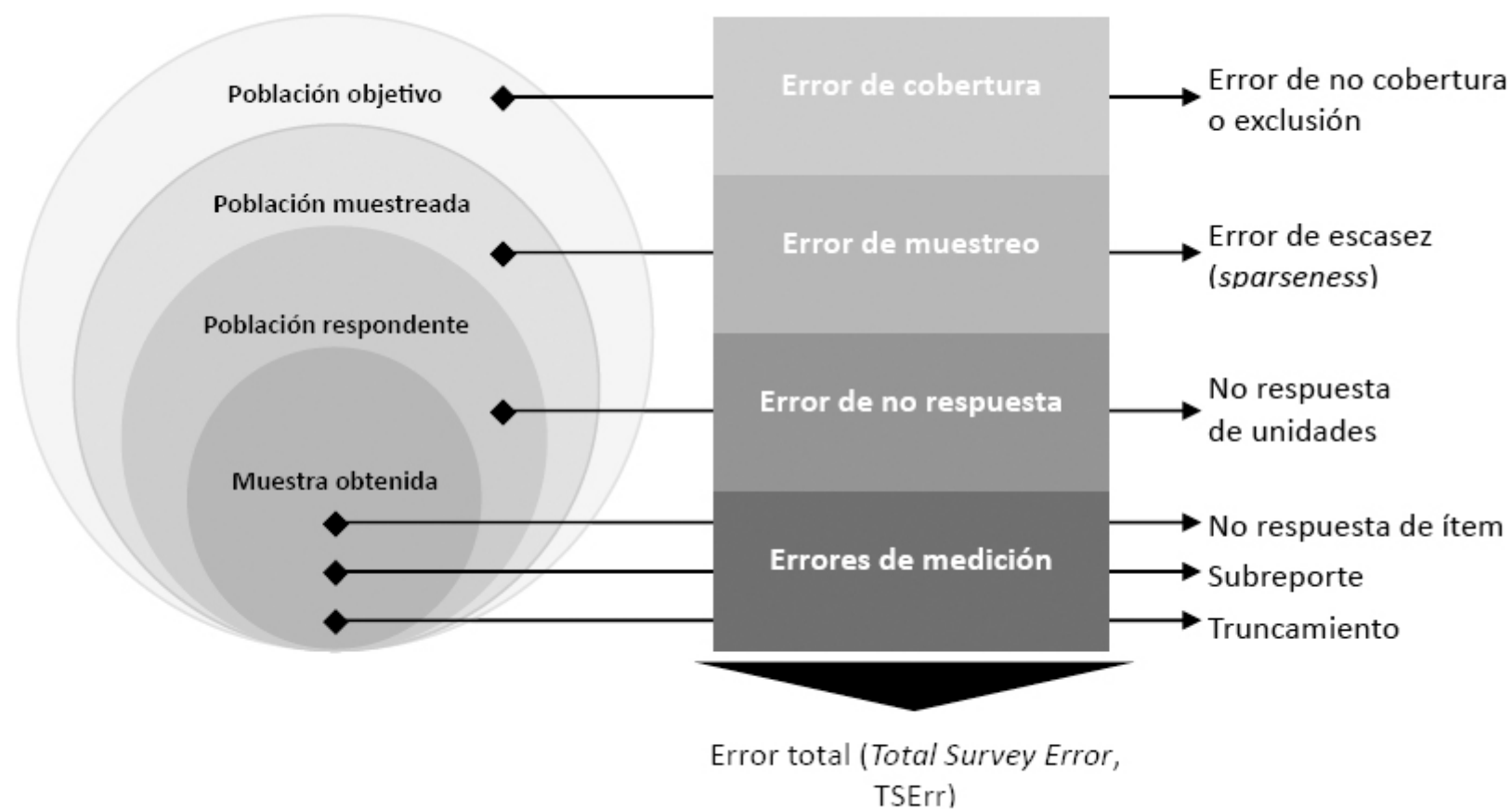
Error de medición no es error muestral

28.2 Marco conceptual

Lograr que los resultados de una encuesta sean representativos de una población determinada implica que cada encuestado esté asociado a un factor de ponderación. Es decir, que cada encuestado tenga un peso o valor que pondere por el número de individuos que representa dentro de la población objetivo. Para ello, la muestra no solo tiene que ser suficiente en términos de cobertura y respuesta, sino que debe incluir a los diferentes subgrupos poblacionales. En la realidad, estas condiciones difícilmente se cumplen, lo que resulta en estimaciones sesgadas (Biemer y Christ, 2008).

El sesgo puede provenir de diferentes fuentes de error dentro del diseño de cobertura, muestreo, respuesta y medición, para lo cual es necesario detallar las poblaciones involucradas en el proceso y los errores relacionados a cada una de ellas (Lustig, 2019; Biemer y Christ, 2008).

Figura 28.1. Correspondencia entre tipos de poblaciones y errores.



Fuente: Elaboración propia con base en la adaptación de Biemer y Christ (2008) y Lustig (2019).

¿Por qué la diferencia entre instrumentos es superior a la diferencia entre periodos?

Ingreso corriente *promedio trimestral*, según año de levantamiento. (Pesos)

Composición de las principales fuentes del ingreso	Promedios (pesos)			
	ENIGH 2016	ENIGH 2018	ENIGH 2020	ENIGH 2022
Ingreso corriente	63 565	60 916	57 370	63 695
Ingreso del trabajo	40 840	40 997	36 612	41 860
Remuneraciones por trabajo subordinado	34 705	34 695	30 936	35 259
Ingresos por trabajo independiente	4 580	4 797	4 305	5 416
Ingresos de otros trabajos	1 555	1 505	1 372	1 185
Renta de la propiedad	5 626	3 591	3 101	3 312
Ingresos provenientes de cooperativas, sociedades y empresas que funcionan como sociedades	4 883	2 863	2 478	2 648
Arrendamientos de activos tangibles y financieros	743	728	623	663
Transferencias	9 871	9 340	10 116	10 928
Jubilaciones y pensiones	4 237	4 244	4 886	5 169
Becas provenientes del gobierno y de instituciones	194	142	139	121
Donativos en dinero provenientes de instituciones y otros hogares	1 335	1 327	1 239	1 271
Ingresos provenientes de otros países	468	481	475	661
Beneficios provenientes de programas gubernamentales	1 120	884	1 330	1 777
Transferencias en especie de otros hogares	1 699	1 547	1 495	1 432
Transferencias en especie de instituciones	819	714	553	499
Estimación del alquiler de la vivienda	7 174	6 934	7 490	7 540
Otros ingresos corrientes	55	54	50	56

Las cifras están expresadas en pesos de 2022.
Nota: La suma de los parciales puede no resultar el total, debido al redondeo.
* Diferencia estadísticamente significativa.
Fuente: INEGI. Encuesta Nacional de Ingresos y Gastos de los Hogares 2016, 2018, 2020 y 2022.

Composición de las principales fuentes del ingreso	Ingreso promedio (pesos)
Ingreso corriente promedio.	66 485
Ingreso del trabajo.	42 640
Remuneraciones por trabajo subordinado.	34 885
Ingresos por trabajo independiente.	6 137
Ingresos de otros trabajos.	1 618
Transferencias.	12 828
Jubilaciones y pensiones.	5 614
Becas provenientes del gobierno y de instituciones.	111
Donativos en dinero provenientes de instituciones y otros hogares.	1 405
Ingresos provenientes de otros países.	683
Beneficios provenientes de programas gubernamentales.	1 906
Transferencias en especie de otros hogares.	2 127
Transferencias en especie de instituciones.	983
Estimación del alquiler de la vivienda.	7 426
Renta de la propiedad.	3 526
Ingresos provenientes de cooperativas, sociedades y empresas que funcionan como sociedades.	2 927
Arrendamientos de activos tangibles y financieros.	600
Otros ingresos corrientes.	64

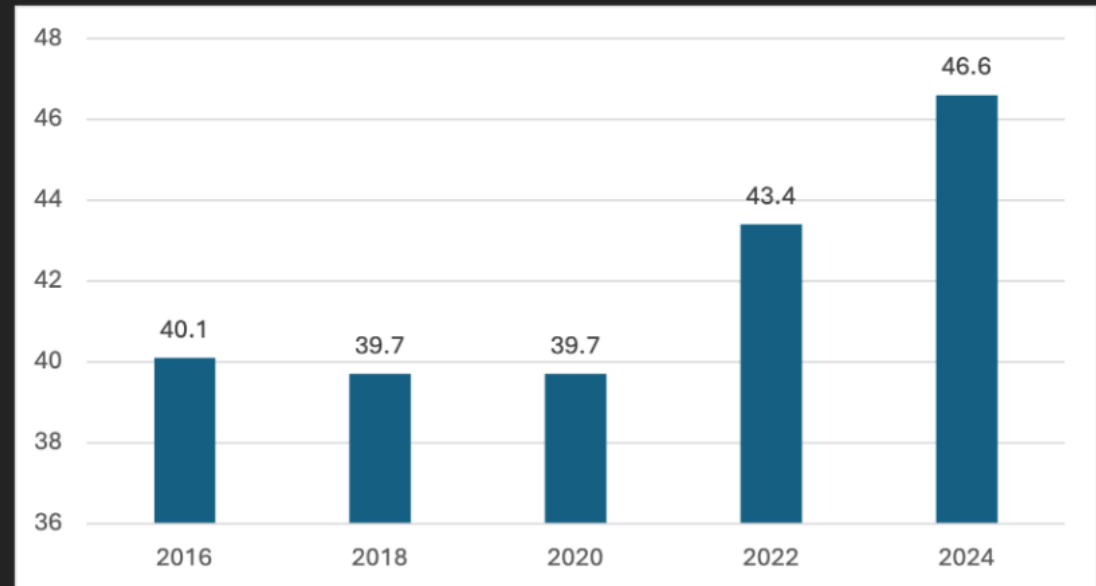
suma de los parciales puede no resultar el total, debido al redondeo.
INEGI. Encuesta Nacional de Ingresos y Gastos de los Hogares Estacional 2022.

La doble vida del ingreso y la pobreza en México

Gerardo Leyva Parra

Agosto 26, 2025

Gráfica 1. Cobertura % del ingreso corriente total de la ENIGH en el de Cuentas Nacionales



Fuente: Cálculos propios a partir de la Encuesta Nacional e Ingresos y Gastos de los Hogares (ENIGH) 2016, 2018, 2020, 2022 y 2024 y las Cuentas por Sectores Institucionales 2016, 2018, 2020 y 2022 y Cuentas por Sectores Institucionales Trimestrales de 2024.

Otros ángulos



**Gerardo
Esquivel**

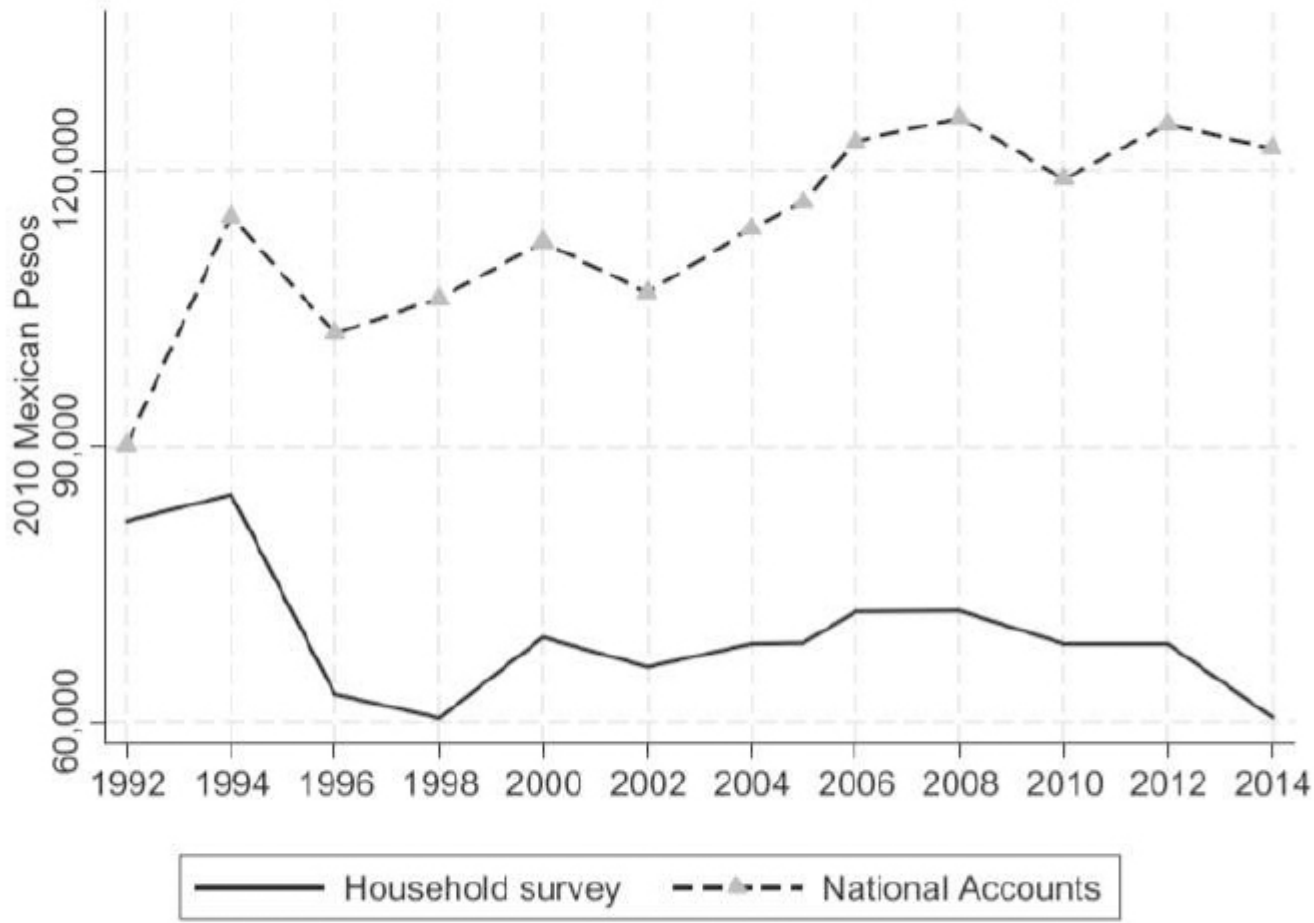
¿Es un mito el aumento del ingreso?

Ahora bien, la pregunta de fondo es si realmente ocurrió el aumento en el ingreso que reportan las personas en la Enigh. Para Leyva, el problema es que en la encuesta las personas reportan un aumento promedio en sus ingresos de 24.2 por ciento real entre 2018 y 2024, lo que parece incompatible con lo reportado en las Cuentas Nacionales (de solo 5.8 por ciento). Esta duda, que pudiera ser legítima, se pudo haber resuelto de manera muy sencilla cotejando otras fuentes que reportan los ingresos de las personas. Antes, sin embargo, vale la pena precisar que el aumento entre 2018 y 2024 en el ingreso promedio de las personas excluyendo transferencias fue de 20.6 por ciento, es decir, ligeramente inferior al del crecimiento de los ingresos totales.

¿Qué dicen las otras fuentes que reportan en México los ingresos de las personas? El salario diario de cotización del IMSS, por ejemplo, creció 22.4 por ciento en términos reales entre 2018 y 2024. Según las Encuestas Mensuales de la Industria Manufacturera (EMIM) y de Empresas Comerciales (EMEC), la remuneración media por persona ocupada en la manufactura creció 21 por ciento en términos reales entre diciembre de 2018 y diciembre de 2024, mientras que en las empresas comerciales creció 26 por ciento. En los tres casos señalados, el rango de

¿Validez
concurrente
o de criterio?

Gráfica 28.1. Ingreso individual anual promedio.



Fuente: Campos, Chávez y Esquivel 2018, Figura 2. ENIGH y cuentas nacionales, INEGI.

Gráfica 28.2. Brecha entre reportes de ingreso ajustados y resultados de encuestas.

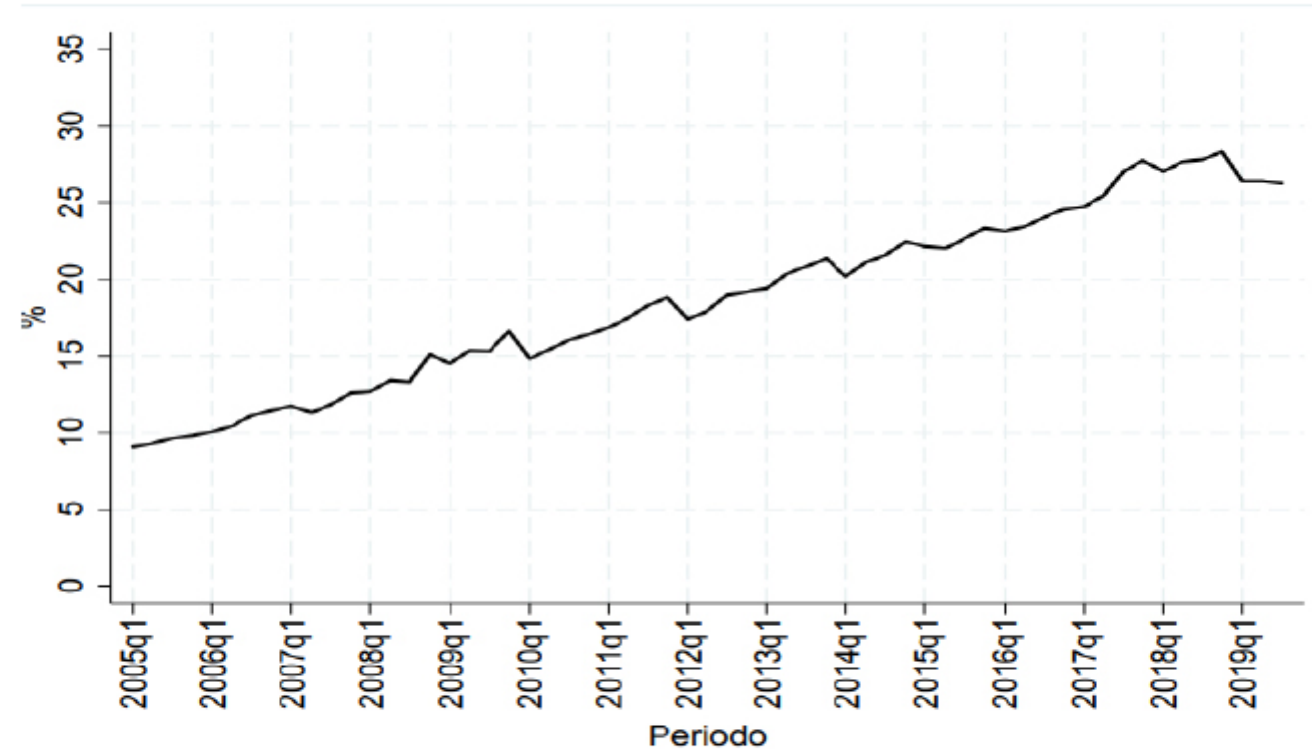
La ENIGH 2024 bajo lupa

Víctor Hugo Pérez Hernández

Septiembre 30, 2025

	2016	2018	2020	2022	2024	Cambio 2018-2024
<i>Pobreza con Ingreso (oficial)</i>						
Pobreza	43.2	41.9	43.9	36.3	29.6	-12.3
Pobreza extrema	7.2	7.0	8.5	7.1	5.3	-1.7
<i>Pobreza con Gasto</i>						
Pobreza	51.4	50.0	52.8	46.7	41.3	-8.7
Pobreza extrema	8.7	8.5	10.8	9.3	7.7	-0.8

Gráfica 28.3. Porcentaje de trabajadores que no declaran ingresos.



Nota: Se graficó para los trimestres 1 a 4. La muestra utilizada se refiere a trabajadores con remuneración, de edad mayor o igual a 12 años, que reportaron haber trabajado un número de horas positivas. No se incluyen trabajadores sin pago.

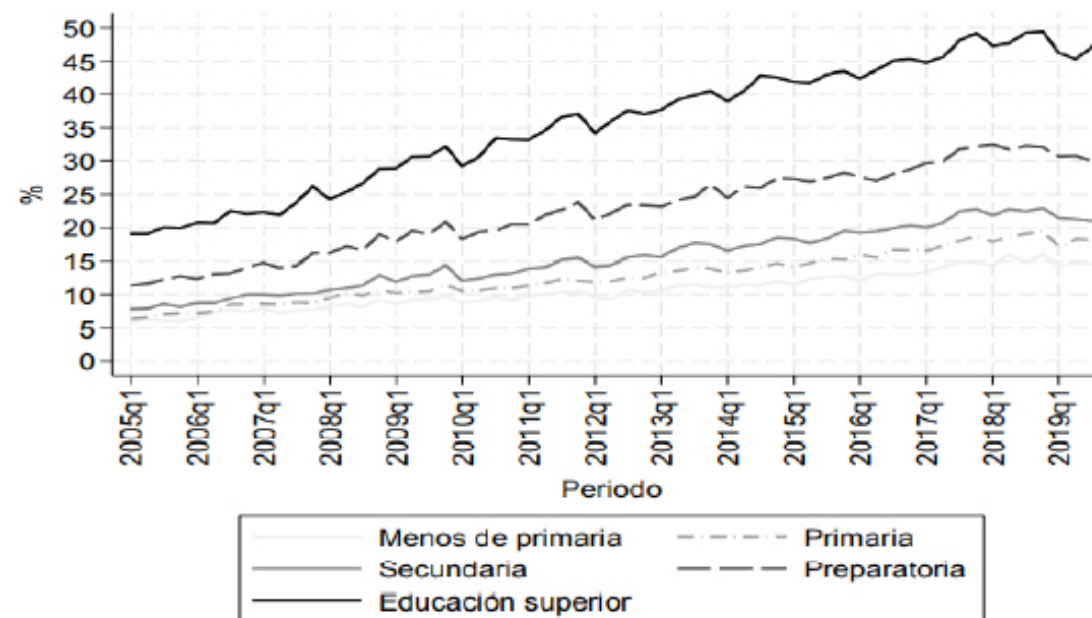
Fuente: Cálculo realizado por el autor con datos de la ENOE.

¿Valores
perdidos
es error de
medición?

¿Error
sistemático?

Gráfica 28.5. Porcentaje de trabajadores que no declaran ingresos, por grupos socioeconómicos.

a) Nivel educativo



¿Error
sistemático?

Statistical Journal of the IAOS 31 (2015) 565–577
DOI 10.3233/SJI-150923
IOS Press

565

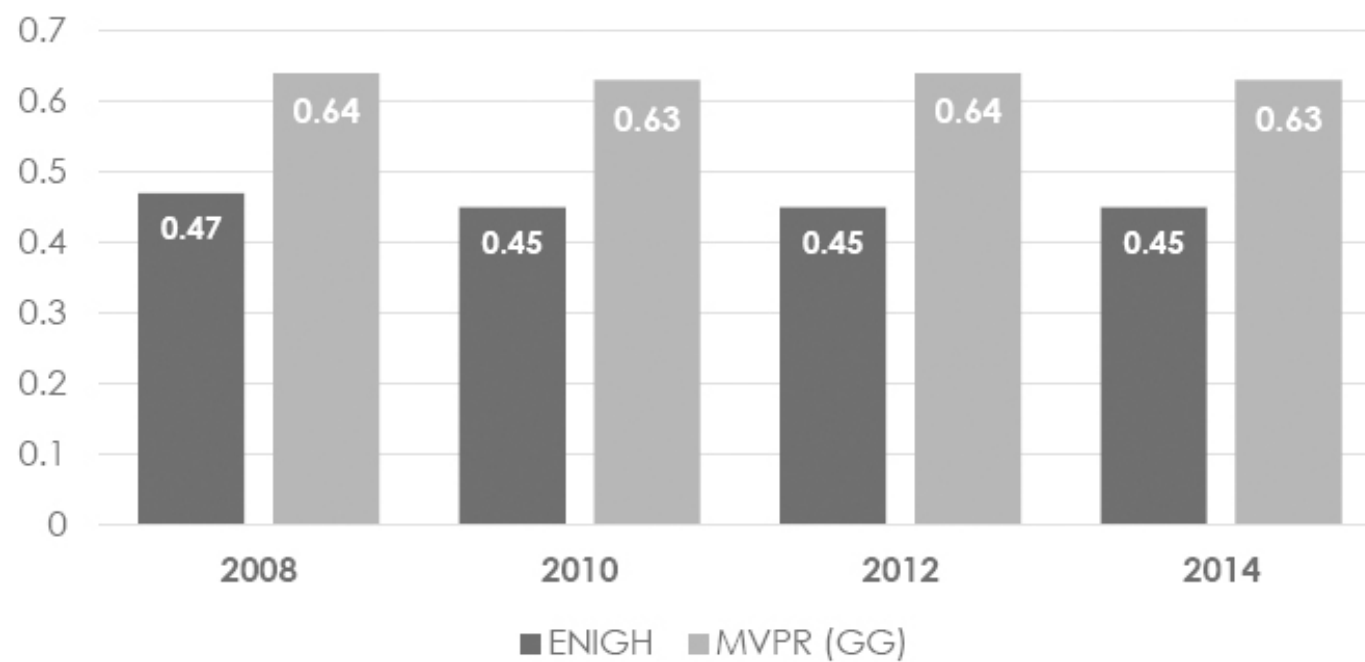
Estimation of the distribution of income from survey data, adjusting for compatibility with other sources

Alfredo Bustos

*Instituto Nacional de Estadística y Geografía (INEGI), H. de Nacozari 2301, Aguascalientes 20276, México
Tel.: +52 449 910 5431; E-mail: alfredo.bustos@inegi.org.mx*

¿Error
sistemático?

Gráfica 28.6. Coeficiente de Gini para México, según diferentes fuentes.

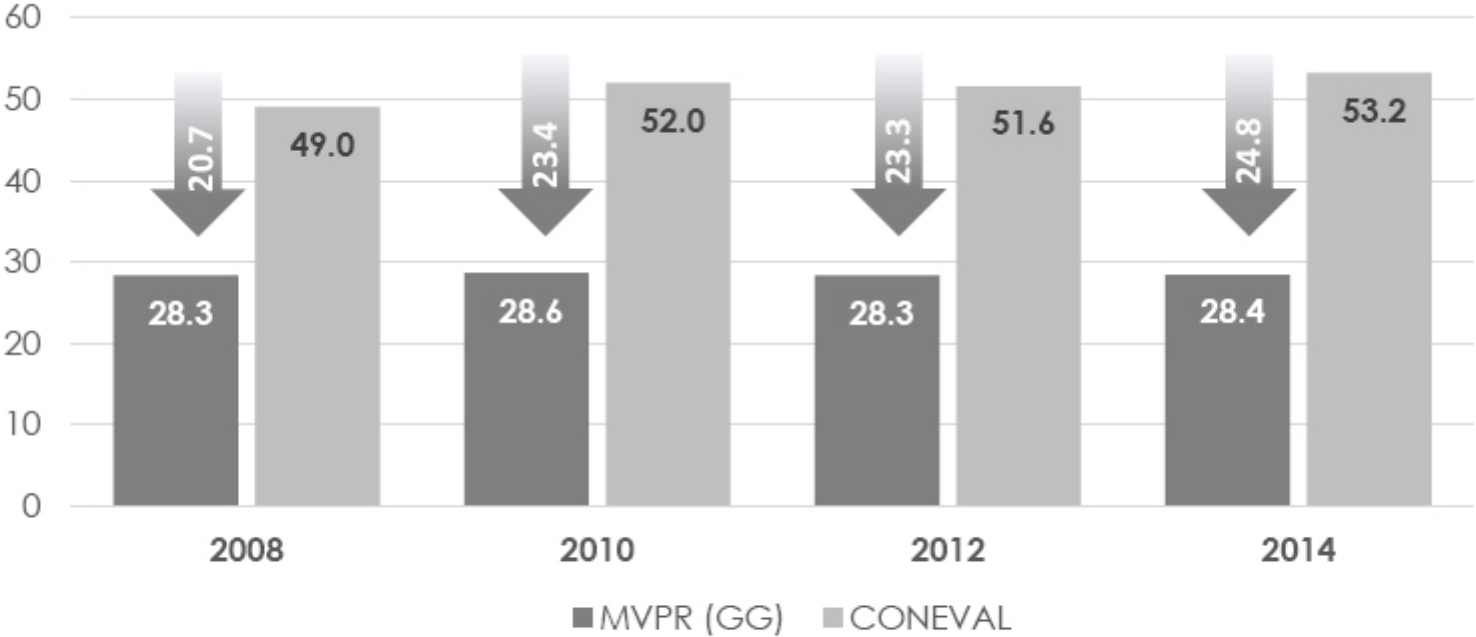


Notas: MPVR se refiere al modelo de Máxima Pseudo-Verosimilitud Restringida utilizado por los autores.

Fuente: Figura 2 de Santaella, Leyva Bustos (2017) con información del INEGI: ENIGH; SCNM y SAT.

¿Error
sistemático?

¿Si el ingreso
está
subestimado la
pobreza está
sobre-estimada?



Notas: MPVR se refiere al modelo de Máxima Pseudo-Verosimilitud Restringida utilizado por los autores. Las barras más claras se estiman con base en la metodología del CONEVAL para la ENIGH.

Fuente: Santaella, Bustos y Leyva (2017) con información del INEGI: ENIGH, SCNM y SAT.

¿Es entonces la imputación la solución al error de medición?

Recomendaciones

Se proponen dos aspectos generales. Por un lado, no solo se tiene que mejorar el diseño y el cuestionario de encuestas, sino que se tienen que usar procedimientos estadísticos para acercar la distribución reportada a la verdadera. Por otro lado, se requiere también usar datos no convencionales como datos georreferenciados e imágenes satelitales para obtener una medida de bienestar diferente a la reportada por el hogar.

En lo que toca al diseño y aplicación del cuestionario, se requiere que la encuesta entreviste a cada miembro del hogar para obtener su ingreso. La práctica de entrevistar a una persona en el hogar y que esa persona reporte el ingreso de cada miembro puede llevar a datos faltantes o mal reportados. Claro que esto conlleva un incremento en costos de la encuesta y habría que analizar si limitando las entrevistas a dos personas por hogar es suficiente. Pero es muy importante que la ENOE pueda incluir entrevistas diferenciadas dentro del hogar. También se tienen que explorar vías para generar mayor confianza con el entrevistado (visitas anteriores a la aplicación del cuestionario, que los sujetos se familiaricen con el entrevistador antes del cuestionario, etc.) para lograr que la pregunta de ingreso sea contestada. Es igualmente importante prever una sobremuestra de hogares de ingresos altos.

La mayoría de las encuestas revisadas utilizan imputación de algún modo. En algunos casos (como en Canadá) se imputa determinísticamente el ingreso de programas sociales. En México hay desconfianza para hacer eso pues no se sabe si verdaderamente el hogar está recibiendo ese ingreso. Un punto intermedio es mencionar el ingreso que debería recibir, preguntar si lo recibe y, en su caso, luego preguntar por qué no lo recibe. Ambas respuestas son importantes en la formulación de políticas públicas. En el caso de ingreso laboral, todos los casos revisados realizan una imputación de ingreso con datos de la misma encuesta y, generalmente, se incluye una columna que identifica los datos que han sido imputados. Aquí se requiere tener un vasto conjunto de características observables, incluyendo datos que no reporta el hogar como calidad de vivienda, calidad de calle, escolaridad en la colonia, entre otros.

Conclusiones

- La confiabilidad se preocupa por el tamaño del error aleatorio de los scores de una escala. Por tanto, la confiabilidad es el techo que tiene una escala, solamente puede empeorarse.
- La validez se encarga de identificar desviaciones sistemáticas en los scores –factores alternativos conocidos o por conocer-
- Validez es un tema de apilar evidencia a favor de la interpretación de los scores. No hay formas substitutas de validez sino complementarias.
- Validez me permite interpretar los scores de una escala con la intención que me propuse

Referencias

- Bock, D. (1997). A brief history of item theory. *Educational measurement: issues and practice*, 16(4), 21-33.
- Lord, F. M., & Novick, M. R. (1968) Statistical Theories of Mental Test Scores. *Reading, Mass.: Addison-Wesley*.
- Ferguson, G. A. (1942). Item selection by the constant process. *Psychometrika*, 7(1), 19-29.
- Finney, D. J. (1944). The application of probit analysis to the results of mental tests. *Psychometrika*, 9(1), 31-39.
- Goldstein, H., & Wood, R. (1989). Five decades of item response modelling. *British Journal. Of Mathematical and Statistical Psychology*, 42, 139–167.
- Goldstein, H. (2012). Francis Galton, measurement, psychometrics and social progress. *Assessment in Education: Principles, Policy & Practice*, 19, 147–158.
- Lawley, D. N. (1943). XXIII.—On problems connected with item selection and test construction. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 61(3), 273-287.
- Markon, K. E. (2018). Reconciling information and reliability in scaling local measurement precision: A comment on O'Connor (2017).
- Nájera Catalán, H. E., & Gordon, D. (2020). The importance of reliability and construct validity in multidimensional poverty measurement: An illustration using the Multidimensional Poverty Index for Latin America (MPI-LA). *The Journal of Development Studies*, 56(9), 1763-1783.
- O'Connor, B. P. (2018). An illustration of the effects of fluctuations in test information on measurement error, the attenuation of effect sizes, and diagnostic reliability. *Psychological assessment*, 30(8), 991.
- O'Connor, B. P. (2018). Clarifications regarding test information and reliability, and new methods for estimating attenuation due to measurement error: Reply to Markon (2018).
- Thurstone, L. L. (1925). A method of scaling psychological and educational tests. *Journal of educational psychology*, 16(7), 433.