Clase 9
¿De qué hablamos cuando hablamos de medir?

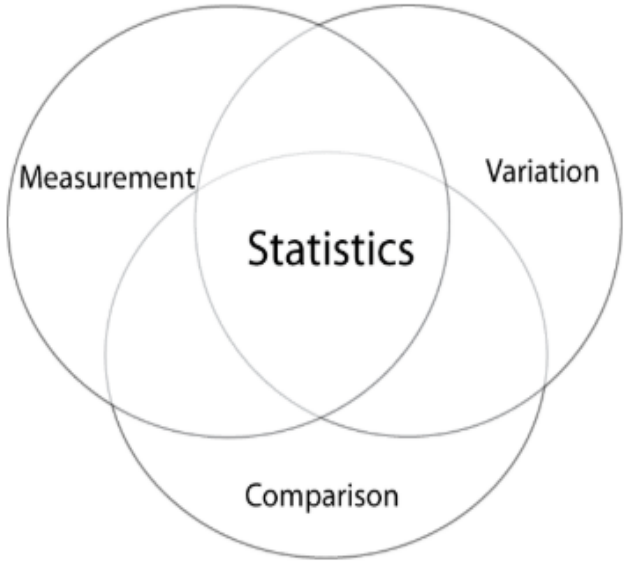
La sesión de hoy

- ¿Por qué debería importarnos la disciplina de la medición?
- ¿Qué es medición?
- ¿Qué es error de medición?
- Medición basada en modelos

Medición

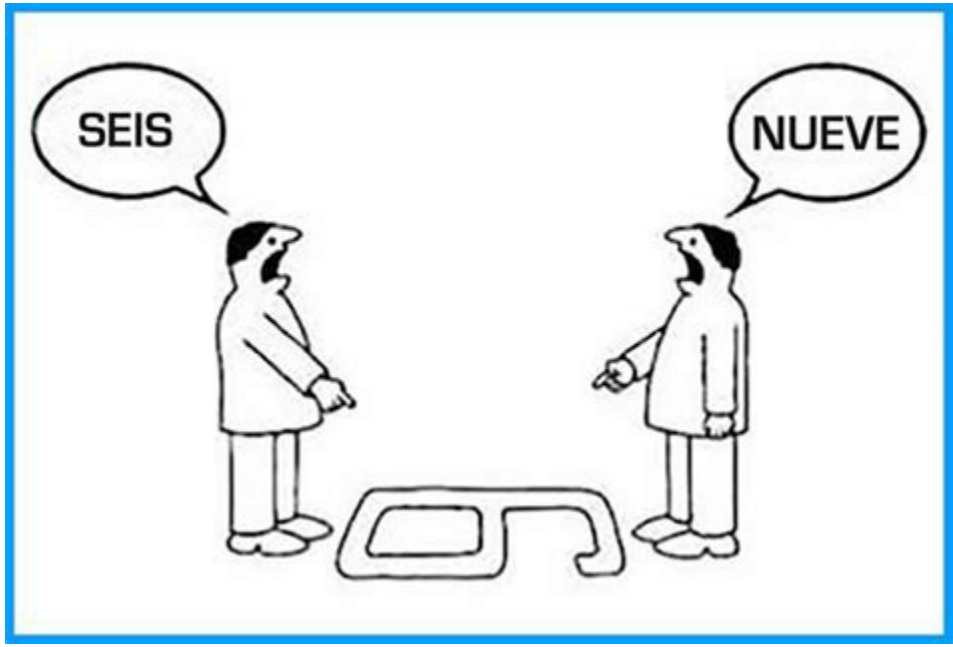
What's the most important thing in statistics that's not in the textbooks?

Posted on April 28, 2015 9:05 AM by Andrew



¿Por qué nos interesa cuantificar?

- Ponernos de acuerdo sobre el tamaño (magnitud)
- Hacer comparaciones inter-temporales (subió / bajo)
- Hacer comparaciones inter-personales (mejor / peor)



¿Qué es medir?

Definición de primer orden:

Generalizar de un conjunto de observables al constructo de interés

¿Por qué medimos? 2

1. Tener más precisión en los juicios que hacemos sobre el mundo
2. Ponernos de acuerdo sobre la magnitud de un fenómeno dado
3. Hacer inferencias con cierta precisión sobre los cambios de magnitud
4. Hacer inferencias con cierta precisión sobre las causas de dichos cambios
5. Hacer inferencias con cierta precisión sobre las consecuencias de dichos cambios
6. Asignar recursos, votos, etc. en función de tales inferencias

Medir no es contar

MEDICIÓN

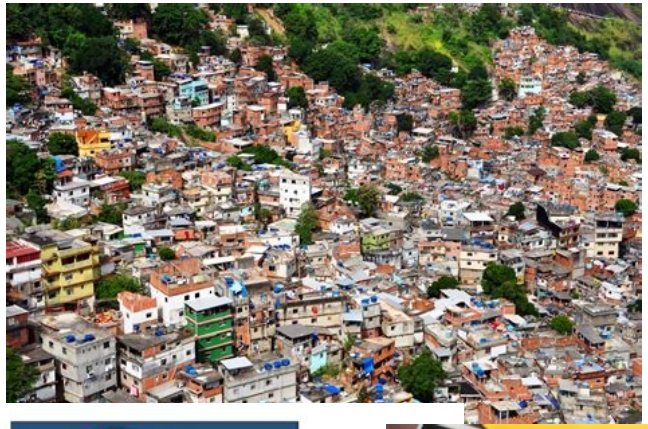
Instrumento, calibración, modelo de medición, agregar, clasificar.
Asignación de numerales a propiedades

$\neq$

CONTAR

Enlistar y en algunos casos agregar

Medir no es contar



INEGI
INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA

CONEVAL
Consejo Nacional de Evaluación de la Política de Desarrollo Social

Cuadro 1
Medición multidimensional de la pobreza*
Porcentaje, número de personas y carencias promedio por indicador de pobreza, 2018 - 2020

Indicadores	Porcentaje		Millones de personas		Carencias promedio	
	2018	2020	2018	2020	2018	2020
Pobreza						
Población en situación de pobreza	41.9	43.9	51.9	55.7	2.3	2.4
Población en situación de pobreza moderada	34.9	35.4	43.2	44.9	2.0	2.1
Población en situación de pobreza extrema	7.0	8.5	8.7	10.8	3.6	3.6
Población vulnerable por carencias sociales	26.4	23.7	32.7	30.0	1.8	1.9
Población vulnerable por ingresos	8.0	8.9	9.9	11.2	0.0	0.0
Población no pobre y no vulnerable	23.7	23.5	29.3	29.8	0.0	0.0
Privación social						
Población con al menos una carencia social	68.3	67.6	84.6	85.7	2.1	2.2
Población con al menos tres carencias sociales	20.2	23.0	25.0	29.2	3.5	3.5
Indicadores de carencia social						
Rezago educativo	19.0	19.2	23.5	24.4	2.8	2.8
Carencia por acceso a los servicios de salud	16.2	28.2	20.1	35.7	2.7	2.8
Carencia por acceso a la seguridad social	53.5	52.0	66.2	66.0	2.3	2.5
Carencia por calidad y espacios de la vivienda	11.0	9.3	13.6	11.8	3.2	3.4
Carencia por acceso a los servicios básicos en la vivienda	19.6	17.9	24.3	22.7	3.0	3.1
Carencia por acceso a la alimentación nutritiva y de calidad	22.2	22.5	27.5	28.6	2.6	2.7
Bienestar económico						
Población con ingreso inferior a la línea de pobreza extrema por ingresos	14.0	17.2	17.3	21.9	2.5	2.5
Población con ingreso inferior a la línea de pobreza por ingresos	49.9	52.8	61.8	66.9	1.9	2.0

*De acuerdo con los *Lineamientos y criterios generales para la definición, identificación y medición de la pobreza* (2018) que se pueden consultar en el Diario Oficial de la Federación (https://www.dof.gob.mx/nota_detalle.php?codigo=5542421&fecha=30/10/2018) y la *Metodología para la medición multidimensional de la pobreza en México,* tercera edición (https://www.coneval.org.mx/InformesPublicaciones/InformesPublicaciones/Documents/Metodologia-medicion-multidimensional-3er-edicion.pdf).
Fuente: estimaciones del CONEVAL con base en la ENIGH 2018 y 2020.

≠

JSTOR

Poverty: An Ordinal Approach to Measurement

Amartya Sen

Econometrica, Vol. 44, No. 2. (Mar., 1976), pp. 219-231.

Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28197603%2944%3A2%3C219%3APAOATM%3E2.0.CO%3B2-Z

Econometrica is currently published by The Econometric Society.

ELSEVIER

Journal of Public Economics
Volume 95, Issues 7–8, August 2011, Pages 476-487

Counting and multidimensional poverty measurement

Sabina Alkire [a, 1], James Foster [a, b, 1]

Show more

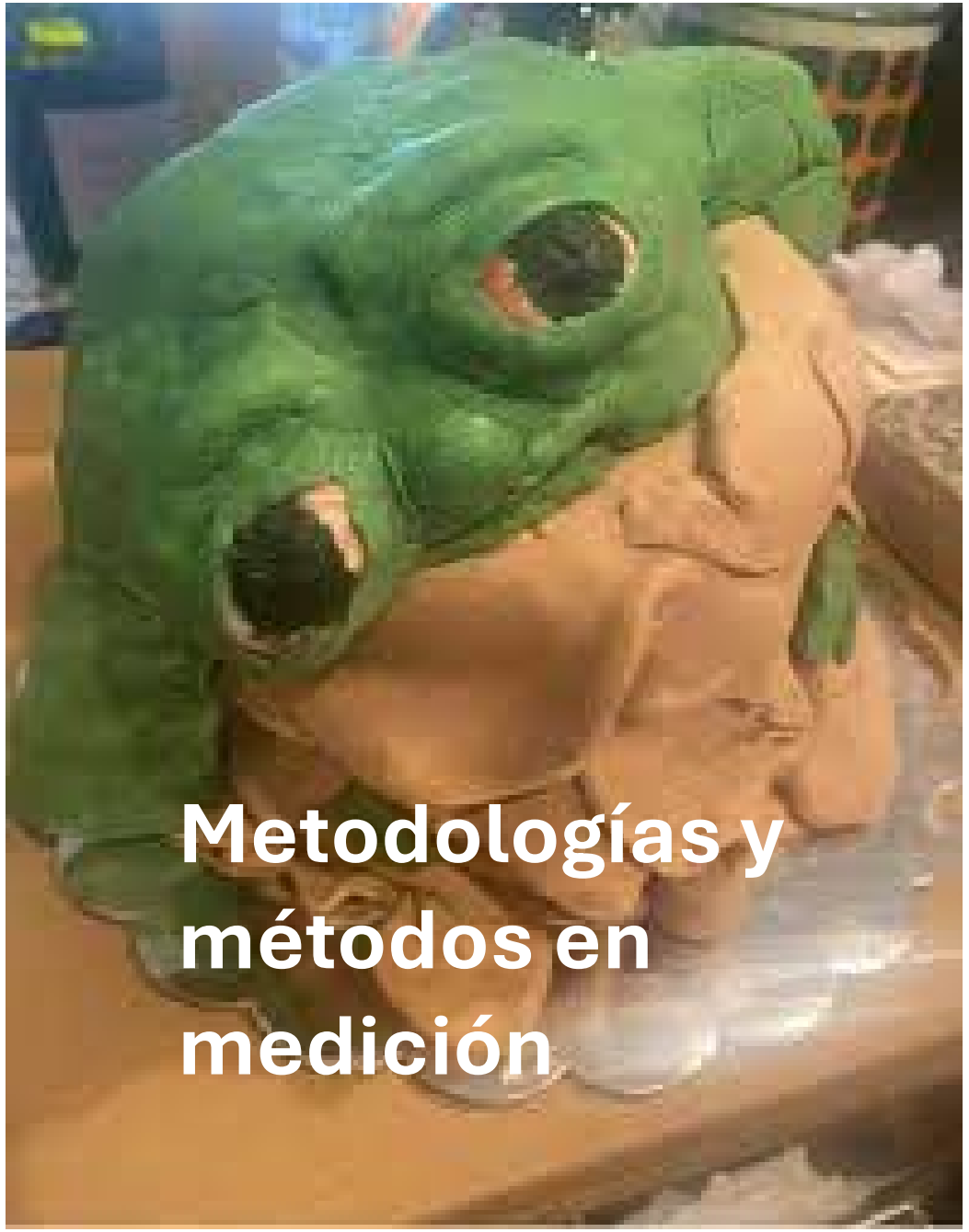
+ Add to Mendeley Share Cite

https://doi.org/10.1016/j.jpubeco.2010.11.006 Get rights and content

Medir es más difícil de lo que nos han hecho creer

- Variación aleatoria:
  - *"Nature is built on unavoidable randomness, limiting what a data-driven view of reality can offer."*
- Validez interna y externa
- Error muestral, error de recolección de datos
- Instrumentos imperfectamente diseñados
- Variabilidad interpersonal
- Variabilidad teórica

Epistemología y
métodos de medición


Metodologías y
métodos en
medición

STATISTICS

Measurement error and the replication crisis

The assumption that measurement error always reduces effect sizes is false

By **Eric Loken**[1] *and* **Andrew Gelman**[2]

Measurement error adds noise to predictions, increases uncertainty in parameter estimates, and makes it more difficult to discover new reliable measurement. In epidemiology, it is textbook knowledge that nondifferential misclassification tends to bias relative risk estimates toward the null (*3*). According to Hausman's "iron law" of econometrics, effect sizes in simple regression models are

Replicabilidad en medición

Si les damos la ENIGH 2020 y les decimos:

Pobreza es falta de recursos en el tiempo y la privación social y material son sus consecuencias

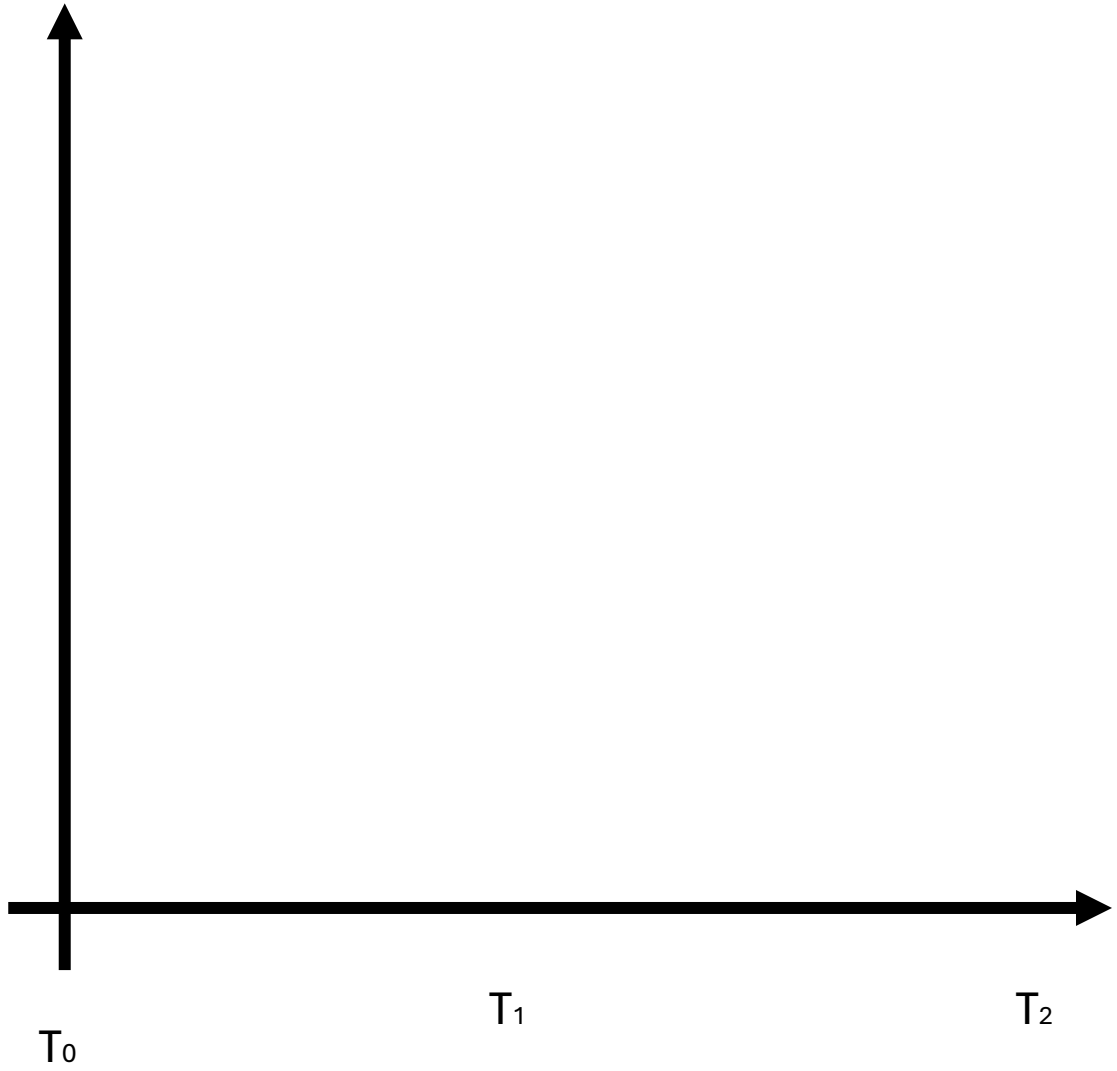
¿Qué tan probable es que todo el grupo obtenga el mismo resultado?

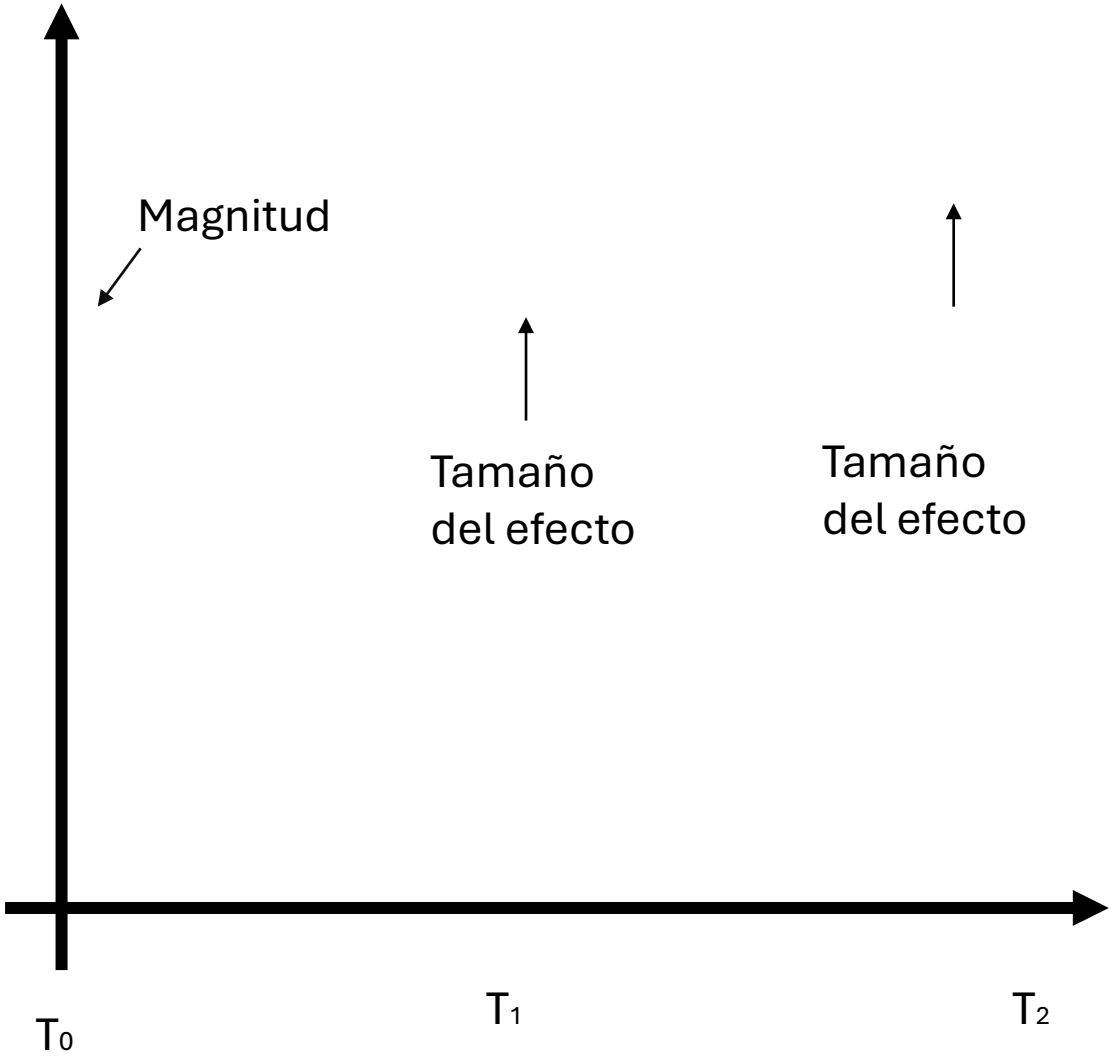
¿Por qué la probabilidad no es 1?

¿Cuáles son las consecuencias?

¿Es inevitable?

Replicabilidad



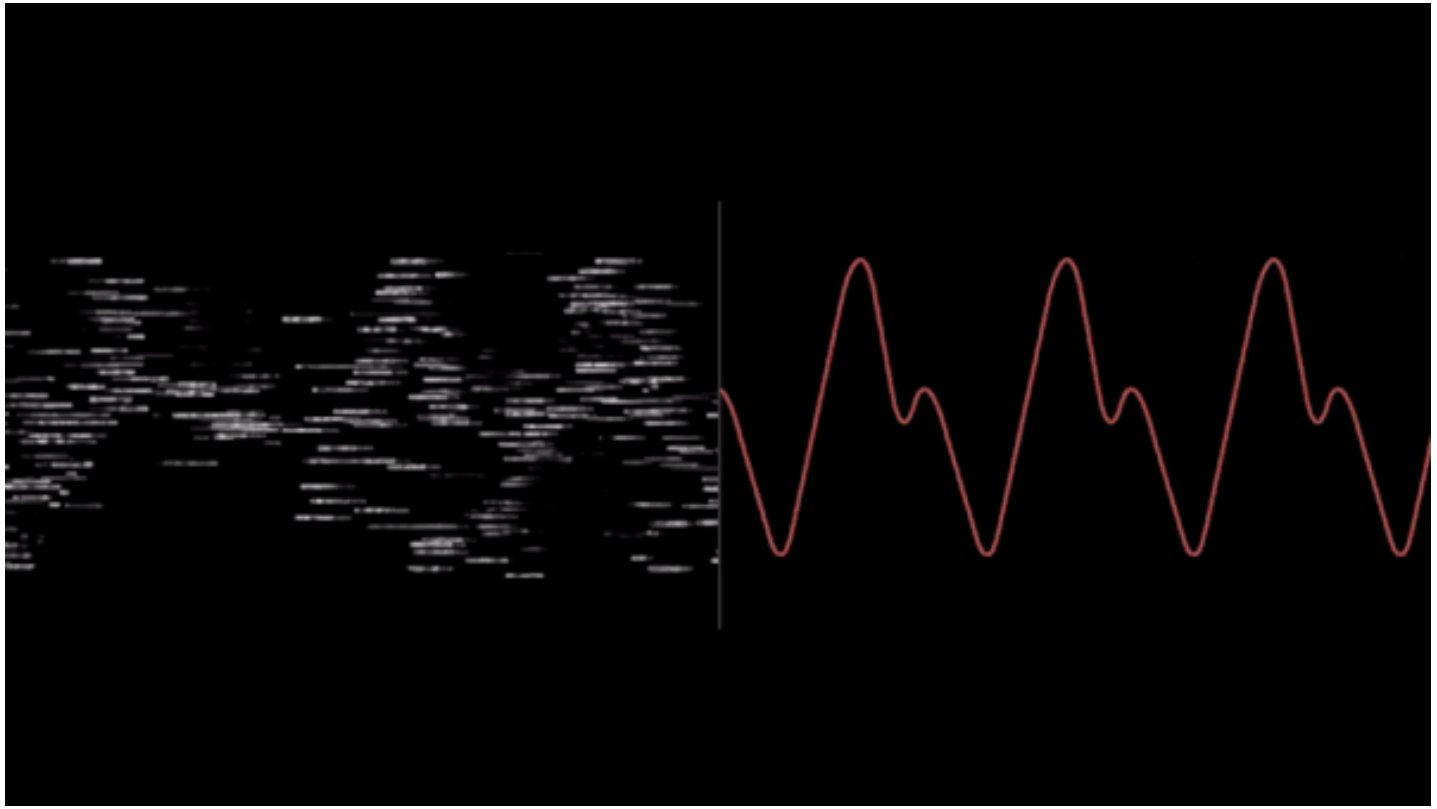


Medición como un tema de incertidumbre / credibilidad

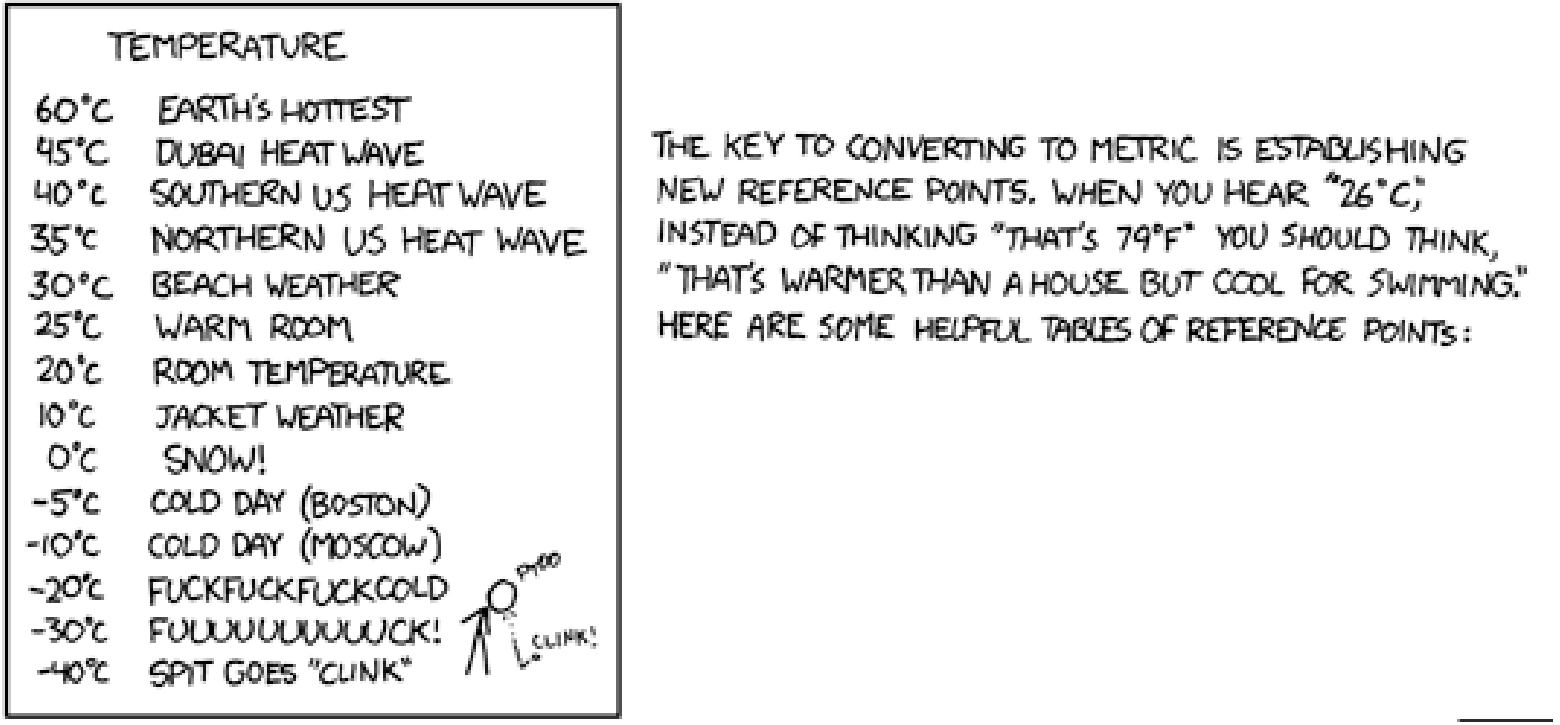
Señal

Entonces la señal es toda aquella variabilidad que nos interesa

En el ideal en medición es cuando logramos capturar con “certidumbre” lo que nos interesa

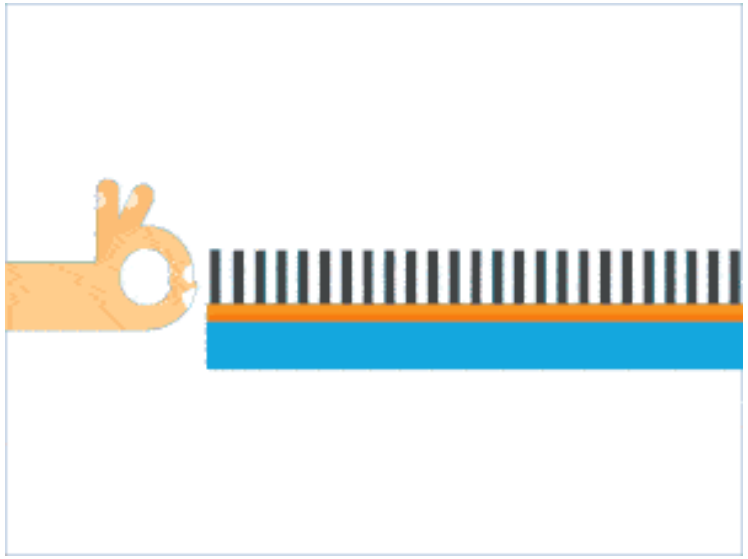


Señal intuitivamente es...



El termómetro captura la suficientemente la variabilidad y los cambios de valor/scores/indicaciones nos hacen sentido

Señal: Causas y consecuencias

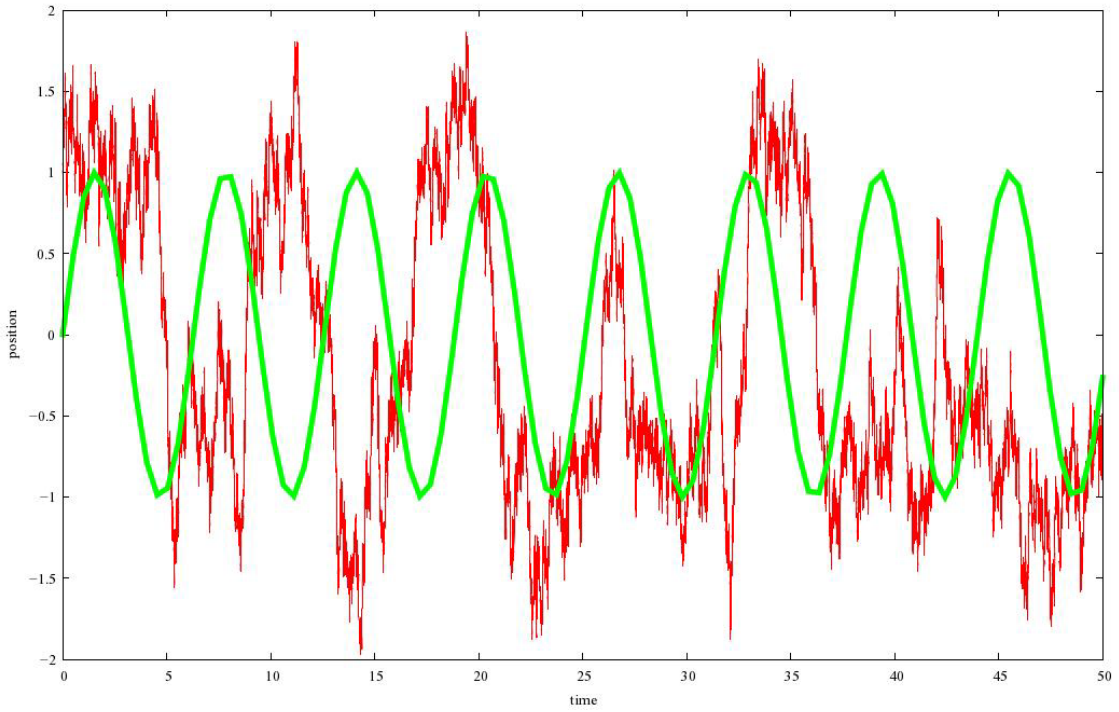


Ruido

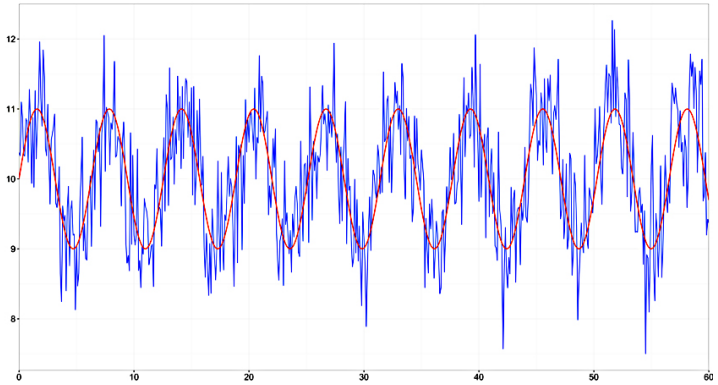
El ruido es toda la variabilidad que NO nos interesa.

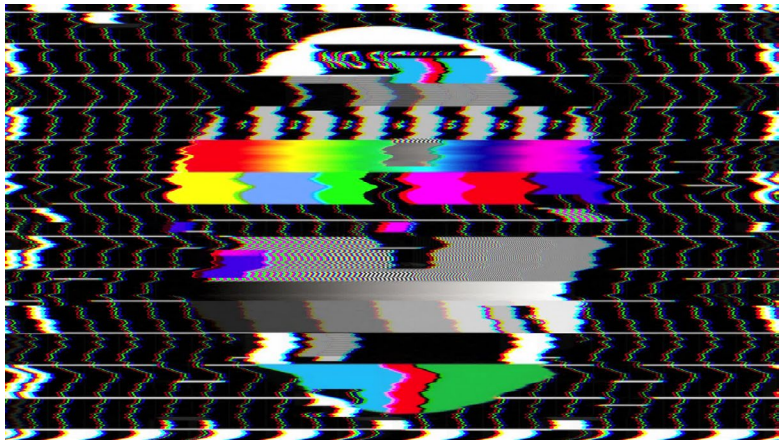
Partamos de que en ciencias sociales medimos fenómenos que varían de persona a persona, son producto de múltiples factores y no son “transparentes” –i.e. variable latente

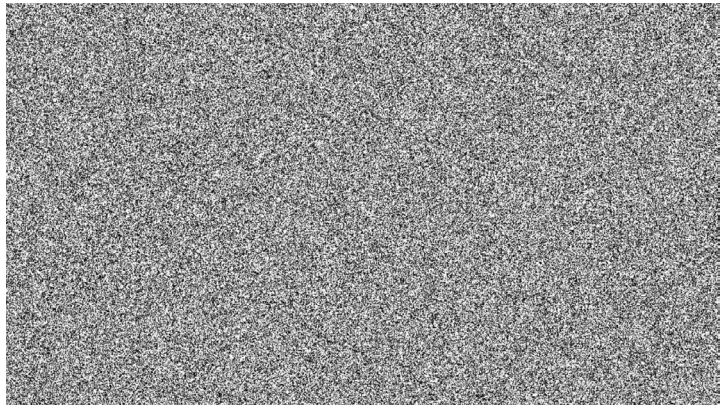
¿De dónde creen proviene el ruido en medición?



El ruido es inevitable





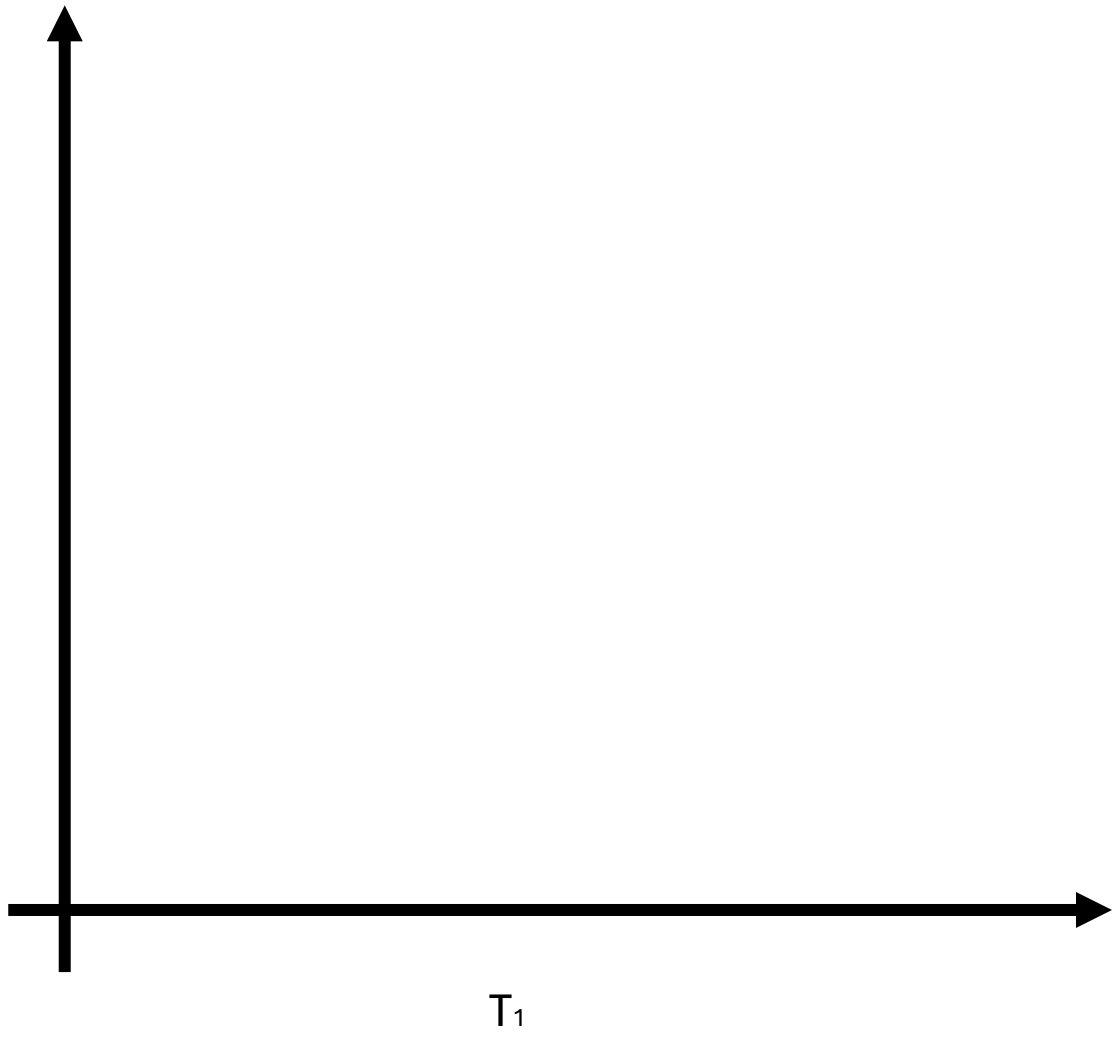


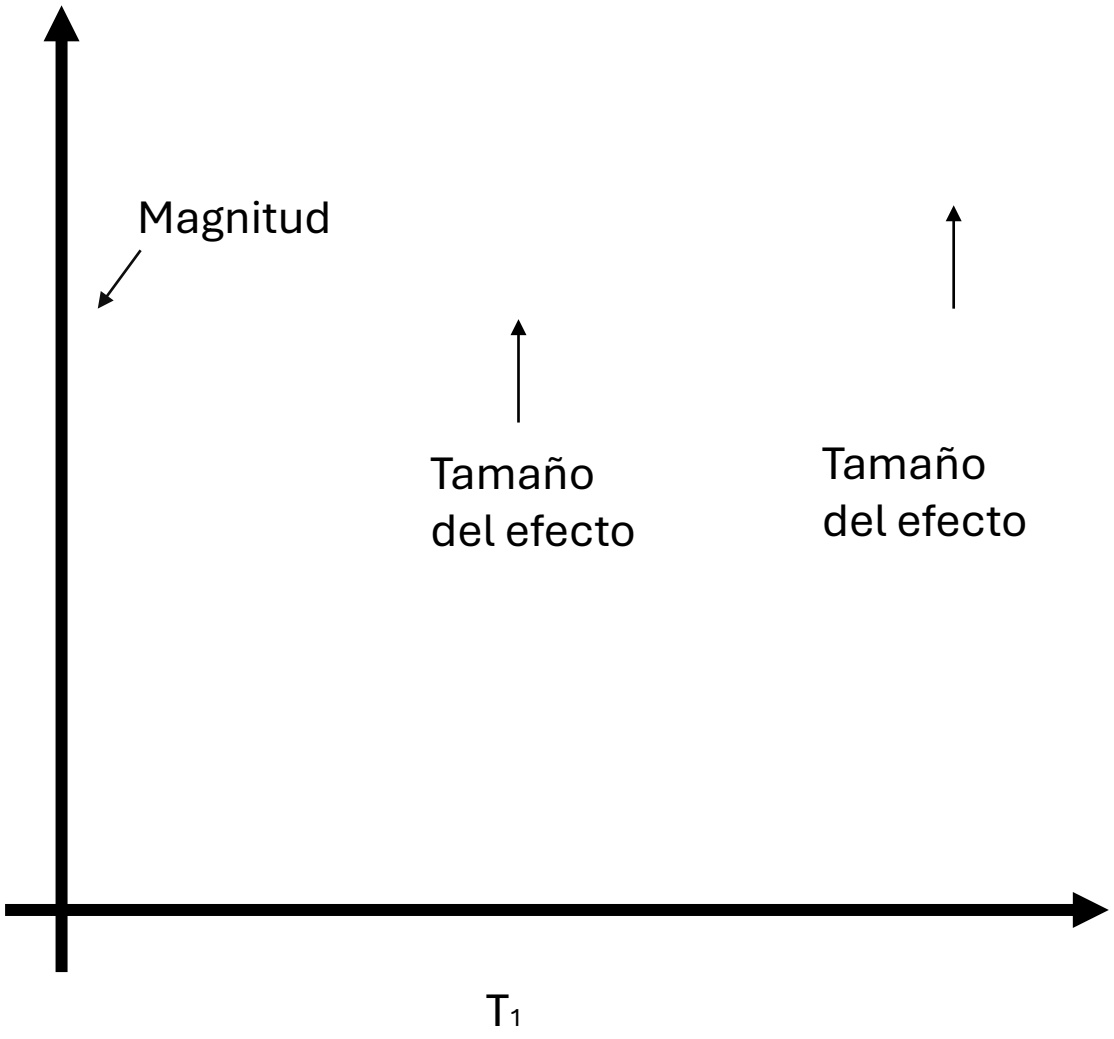
Poco ruido a mucho ruido (error inaceptable)

Error de medición

“Es la distancia entre eso que quería medir y lo que terminé midiendo”

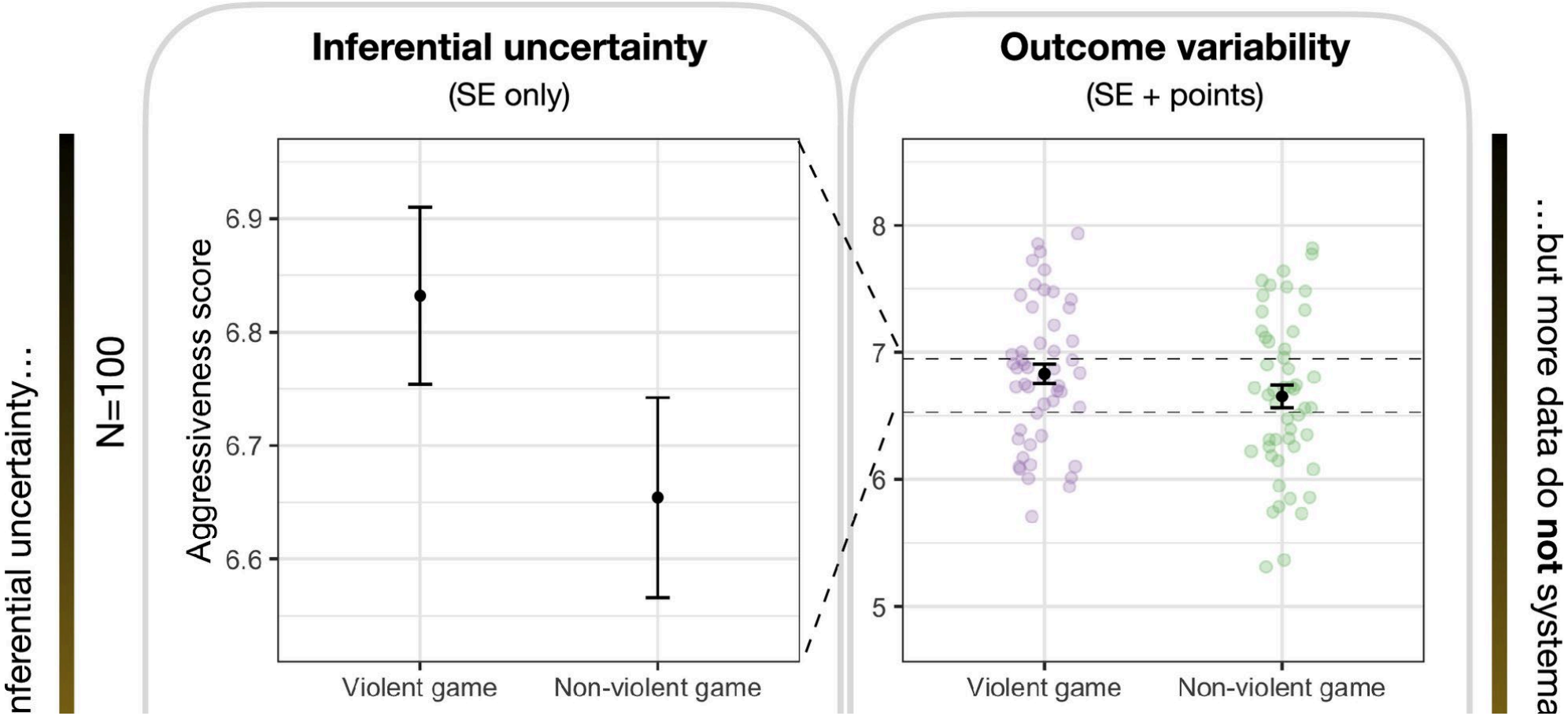
“Measurement error adds noise to predictions, increases uncertainty in parameter estimates, and makes it more difficult to discover new phenomena or to distinguish among competing theories.” Loken y Gelman (2017)



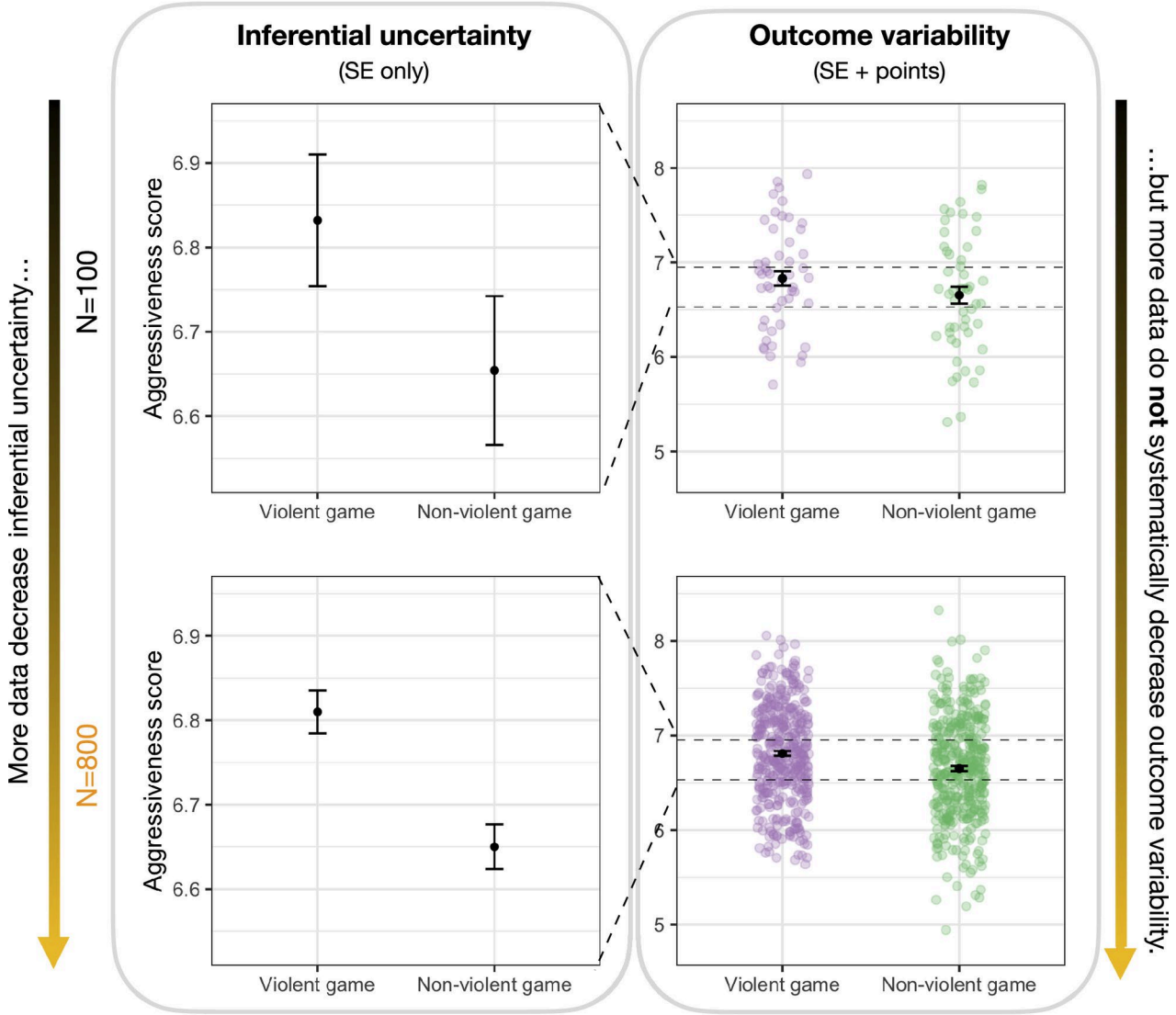


Alto error de medición
Replicabilidad
Everything was fine except. . .
DIVORCE COURT
4GIFs.com
FOX2

Mitos en inferencia y replicabilidad



El error de medición no es el error muestral



Puedes tener resultados significativos con errores de medición

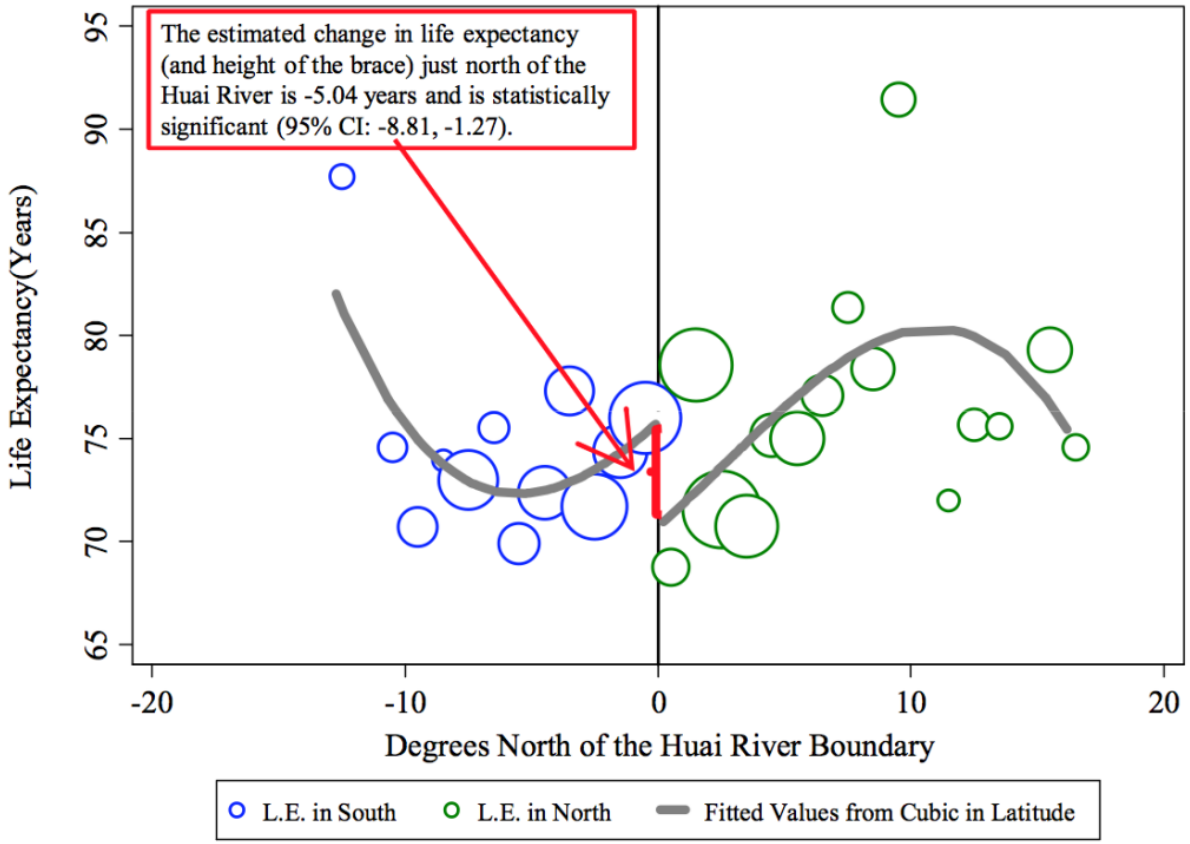
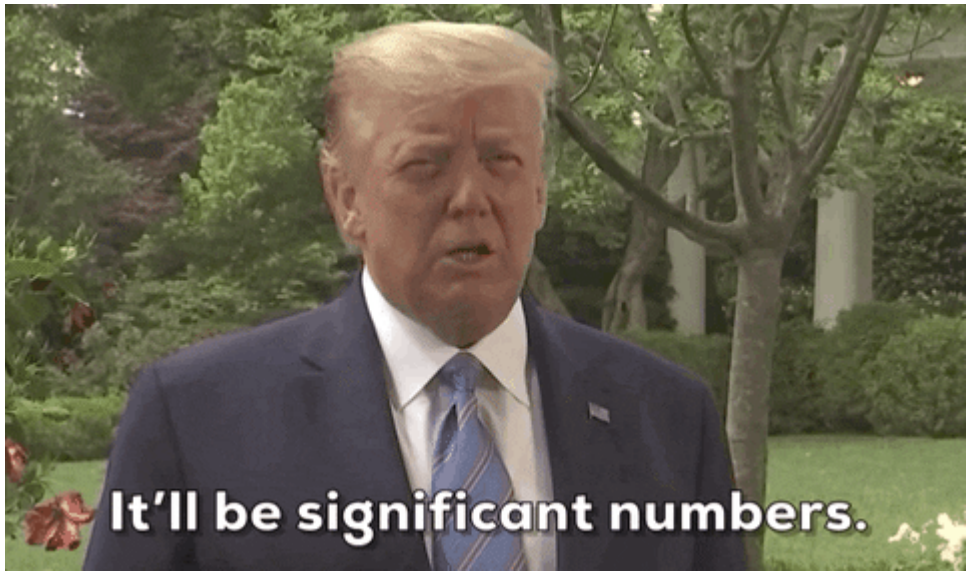


Fig. 3. The plotted line reports the fitted values from a regression of life expectancy on a cubic in latitude using the sample of DSP locations, weighted by the population at each location.

Pero los resultados son significativos...



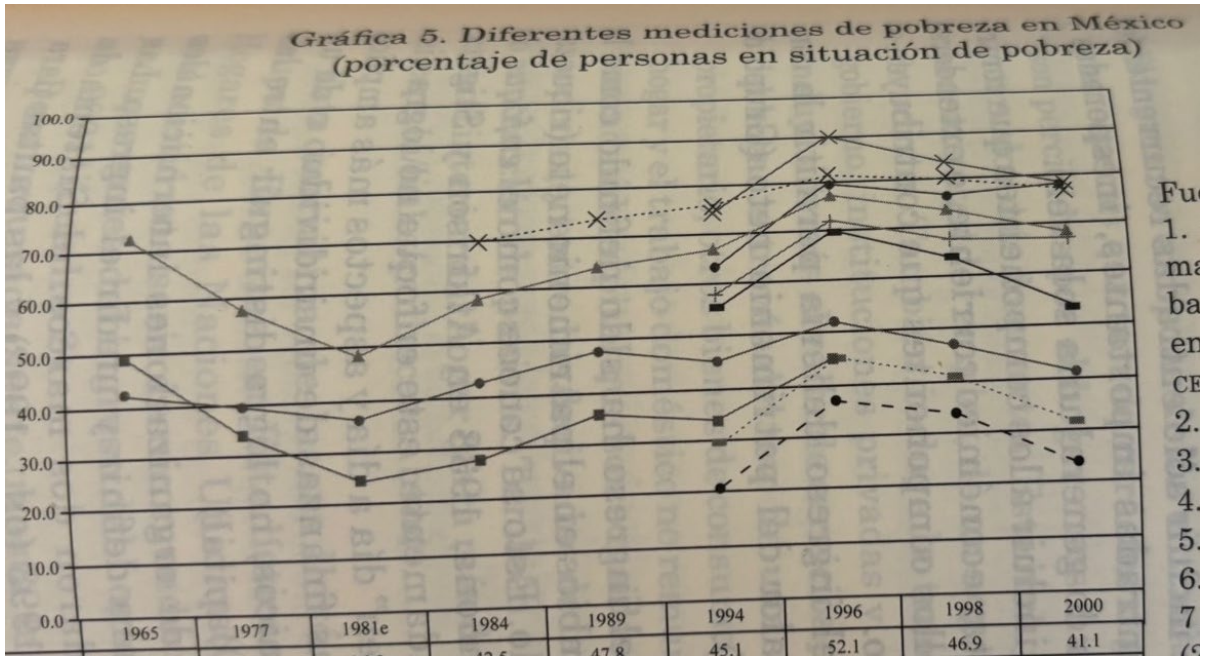
Cinco mitos de la inferencia

1. La medición es un tema exclusivo de varianza y sesgo
2. En estudios controlados el sesgo no importa
3. La varianza importa porque si tu varianza es más alta, tus errores estándar serán más altos y tendrás menos chances de alcanzar significancia estadística
4. Si tengo alta variabilidad, los errores estándar se reducen y por tanto un $p<.05$ es indictivo de éxito
5. Si medí Z en lugar de X, pero lo que me interesaba era la relación entre X y Y, siempre puedo cambiar a la relación de Z y X.

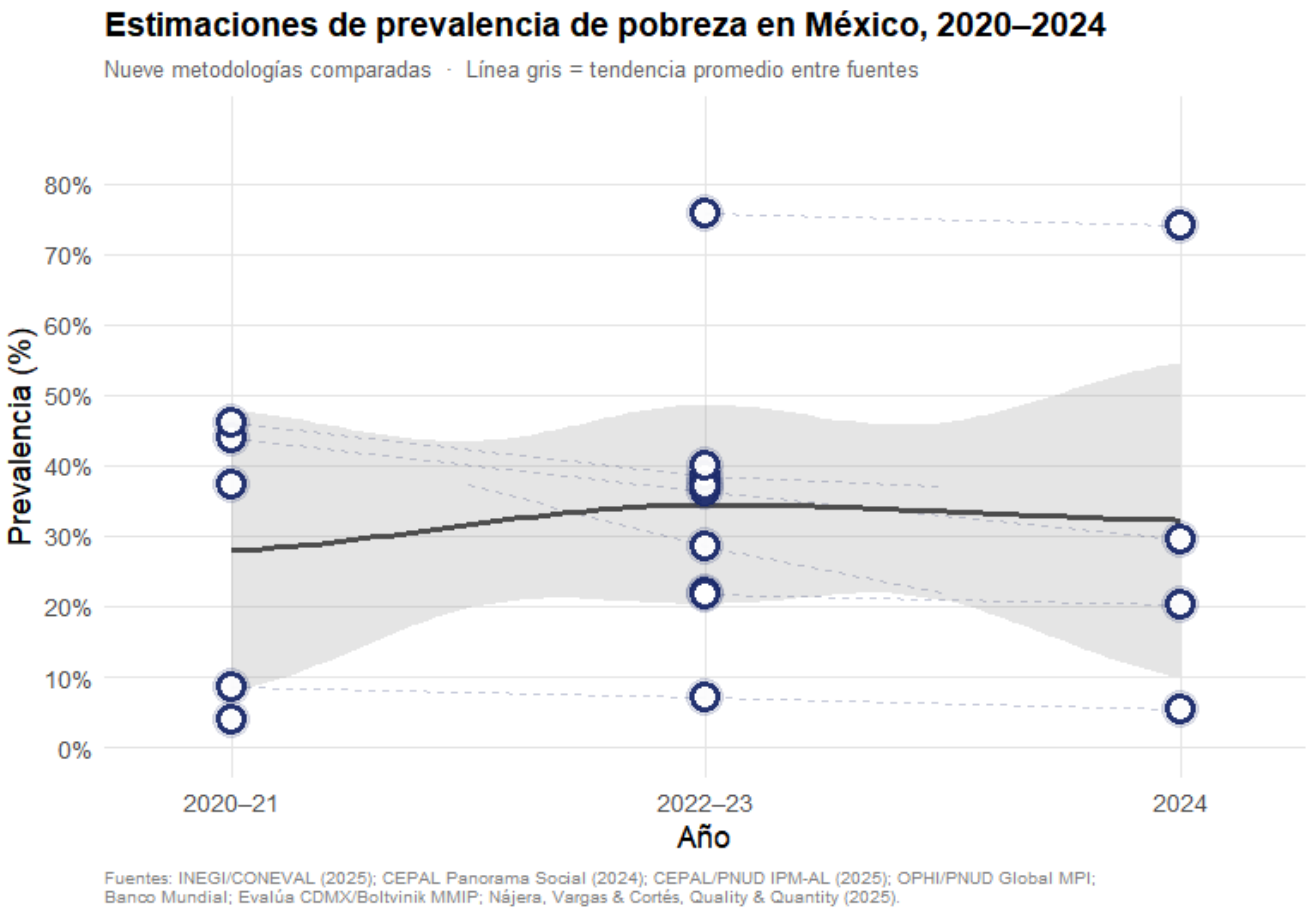
Error de medición

“Es la distancia entre eso que quería medir y lo que terminé midiendo”

Principios de siglo



25 años después



Error de medición como juez

Error de medición

- Para calcularlo, basta entonces conocer el valor de eso que quería medir

Los objetos científicos no están a la mano

- Tiempo
- Temperatura
- Dureza
- Firmeza
- Inflación
- Desempleo
- Informalidad
- Clase Social
- Pobreza
- Discriminación
- Presión atmosférica

Error de medición

Es un tema de modelos **teóricos+estadísticos**:

- Pobre definición del objeto científico
- Problemas en la generación de datos:
  - Uso ad hoc de información disponible
  - Problemas de recolección
  - Problemas de codificación
  - Problemas de muestreo
- Supuestos del modelo teórico que no se cumplen durante el proceso de la medición

¿Cómo detectar la magnitud de dichos problemas?

¿Medición sin teoría?

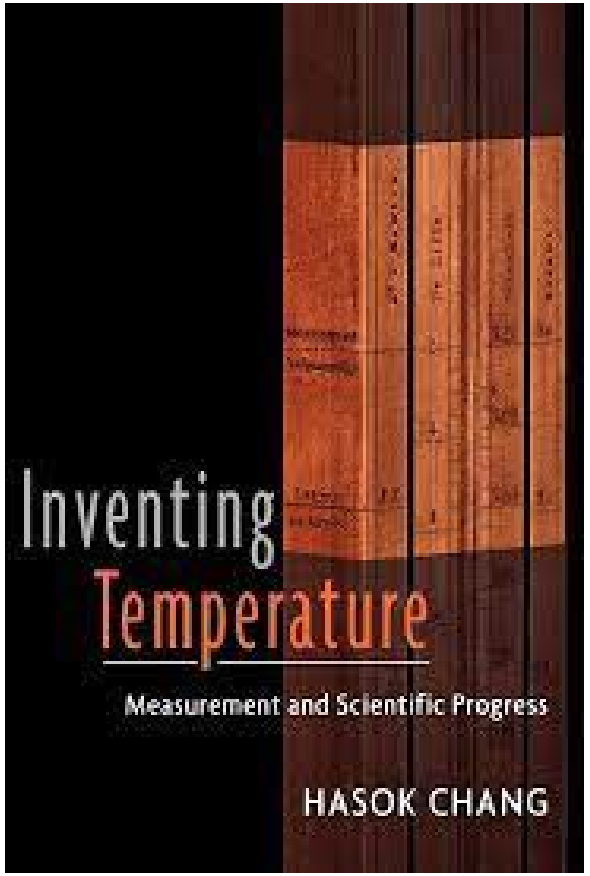
Teoría:

- Concepto
- Definiciones
- Enunciados

SI vamos a medir la temperatura

¿Cómo sabemos sin proposiciones claras de qué significa?

¿Quizá solo podemos medir ciertas cosas?



Brit. J. Phil. Sci. **67** (2016), 297–335

Making Time: A Study in the Epistemology of Measurement

Eran Tal

How Accurate Is the Standard Second?

Author(s): Eran Tal

Source: *Philosophy of Science*, Vol. 78, No. 5 (December 2011), pp. 1082-1096

Published by: The University of Chicago Press on behalf of the Philosophy of Science Association

Stable URL: https://www.jstor.org/stable/10.1086/662268

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide

¿Podemos ignorar estas preguntas?

- ¿Qué es un resultado de medición?
- ¿Qué hace certero/exacto/válido (accurate) a un resultado de medición?
- ¿Qué justifica el uso de resultados de medición como evidencia científica?

Medición basada en modelos

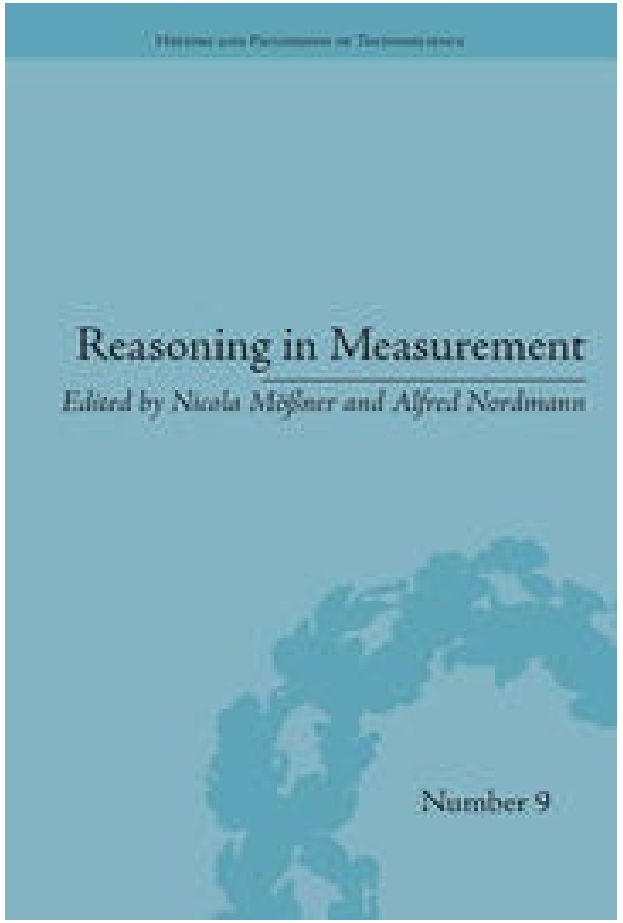
Brit. J. Phil. Sci. **67** (2016), 297–335

Making Time: A Study in the Epistemology of Measurement

Eran Tal

ABSTRACT

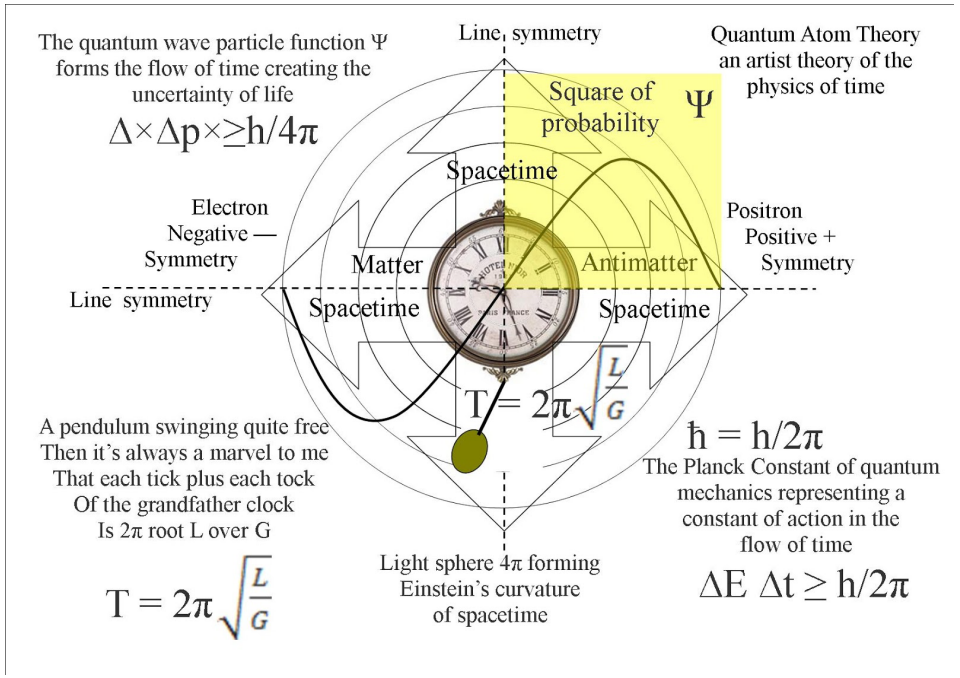
This article develops a model-based account of the standardization of physical measurement, taking the contemporary standardization of time as its central case study. To standardize the measurement of a quantity, I argue, is to legislate the mode of application of a quantity concept to a collection of exemplary artefacts. Legislation involves an iterative exchange between top-down adjustments to theoretical and statistical models regulating the application of a concept, and bottom-up adjustments to material artefacts in light of remaining gaps. The model-based account clarifies the cognitive role of *ad hoc* corrections, arbitrary rules, and seemingly circular inferences involved in contemporary timekeeping, and explains the stability of networks of standards better than its conventionalist and constructivist counterparts.



Modelos de medición

Instrumentos

Modelo teórico y parámetros de la medición



Medición basada en modelos

Modelo teórico

Objeto científico
Propiedades del sistema bajo medición
Supuestos / Proposiciones

+

Modelo estadístico

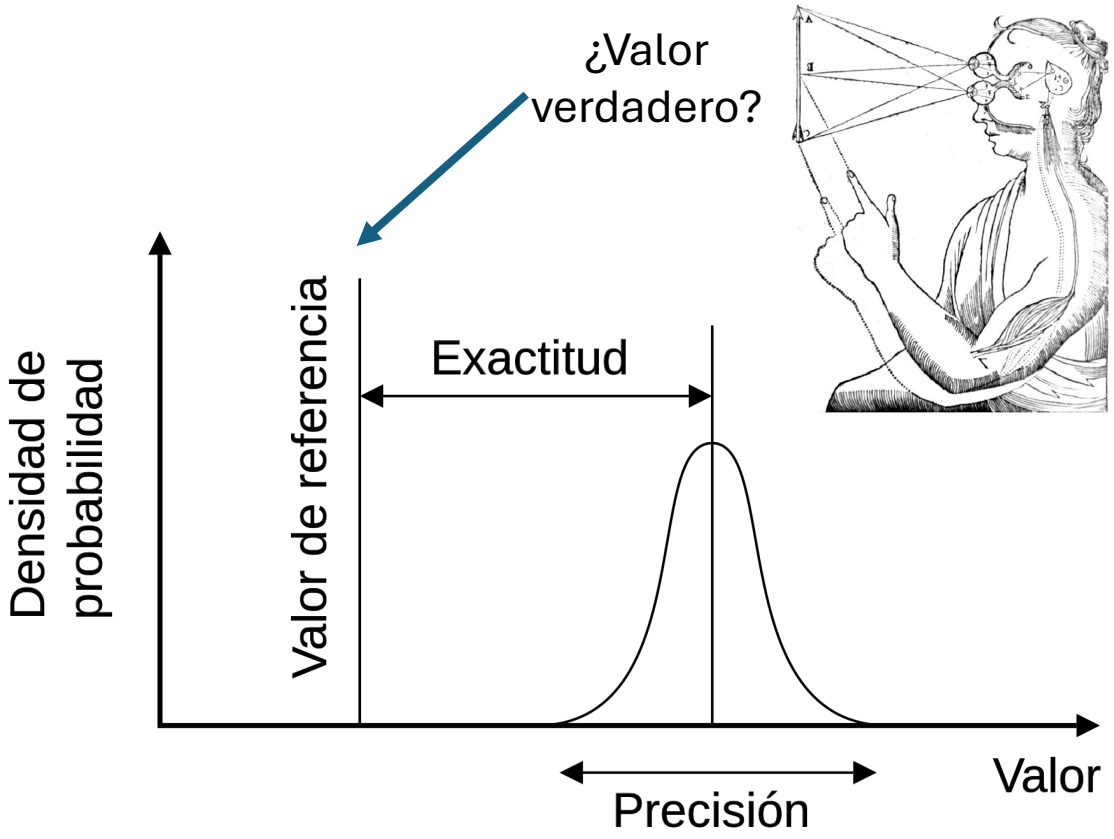
Experimento: parámetros de los supuestos y proposiciones (SEM)

Sin modelo, no hay medición

¿Exactitud (accuracy) y error de medición?

- ¿Cómo evaluar nuestras presunciones de conocimiento?
- ¿Qué confianza dar a nuestros resultados?
- ¿Cómo saber que estamos midiendo lo que nos propusimos medir en primer lugar y no cualquier otra cosa?
- ¿Cómo cuantificar el error involucrado en nuestras medidas?

$$X = True + error$$



Caracterizaciones recientes

- Orientadas a la práctica (¿qué hacen los que miden?)
  - Reconocen la riqueza de los medios representacionales involucrados
  - Particularmente el uso generalizado de supuestos teóricos en
    - el diseño de instrumentos de medición y en
    - la interpretación de sus indicaciones
  - Reciente conjunto de investigaciones (epistemología de la medición)
    - Actividad de recolección de información basada en modelos (del proceso de medición mismo)
    - Metateoría conceptualmente consistente (sin ambigüedades) para enmarcar el panorama del problema

Fenómenos
(ante los ojos)

≠

Objetos científicos

Fenómenos
(ante los ojos)

≠

Observación
(codificada)

Datos

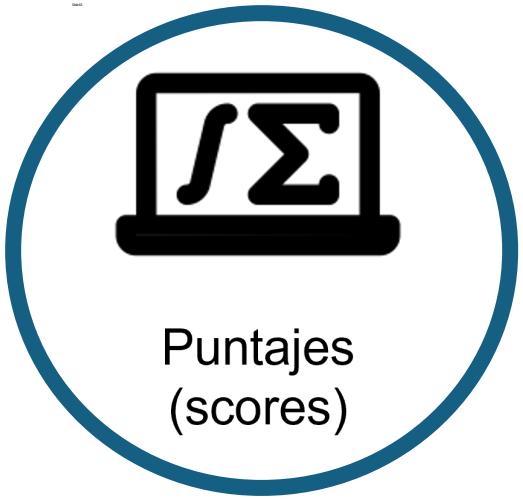
≠

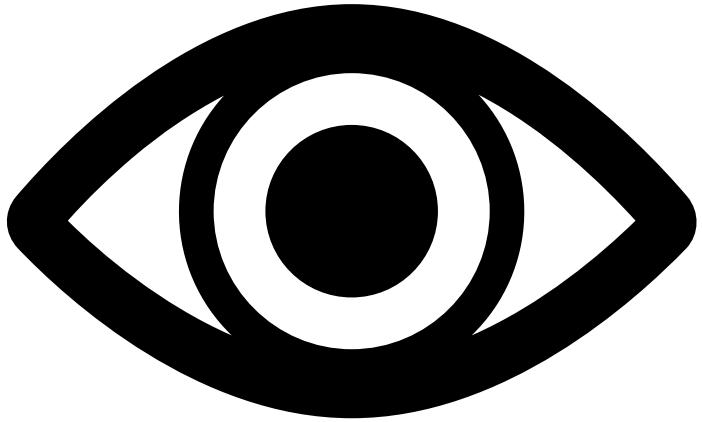
Estimadores

Puntajes

≠

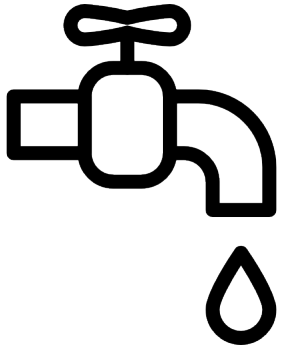
Objetos científicos





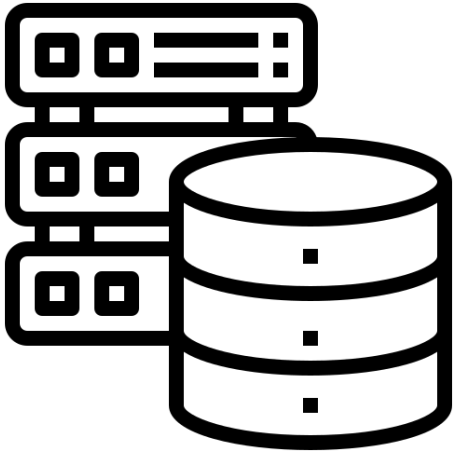
Sistema bajo medición

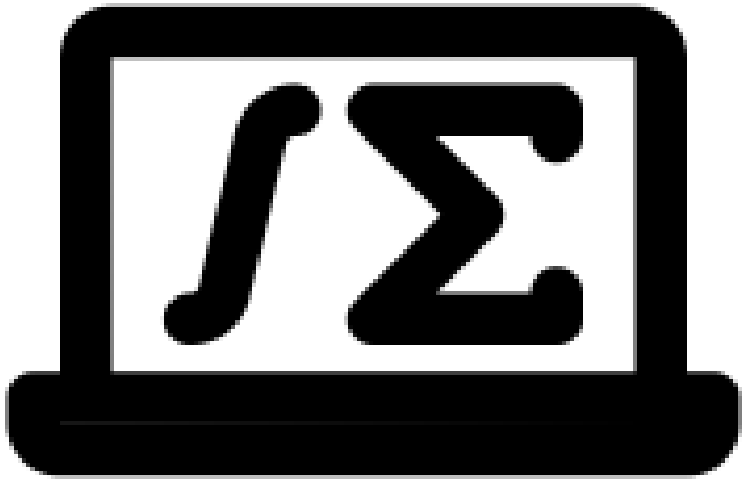
- Los referentes
- El mundo (natural) allá afuera
- Los fenómenos (ante los ojos)



Indicaciones instrumentales

- Lectura/propiedad de los instrumentos
- Generación/fuente de datos
- Indicaciones (sin compromiso)
- Variables en bases



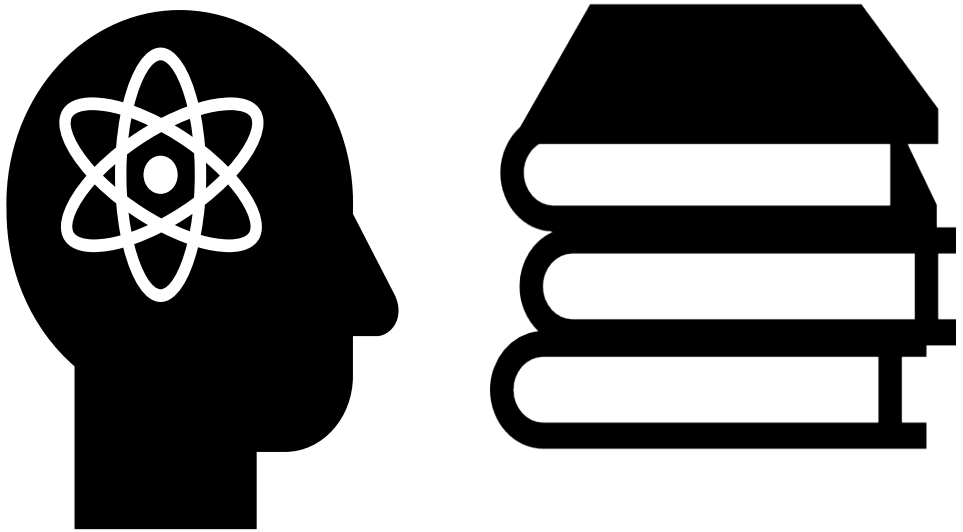


Puntajes (scores)

- Procesamiento/transformación/ajuste de datos
- Modelaje estadístico
- Método de agregación

$$M_0 = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{d} w_j g_{ij}^{0}(k)$$

$$\mathrm{X_{ik}} = [\ \mathrm{a_k} + \mathrm{b_k}\ (\mathrm{T_i})\ ] + \mathrm{E_{ik}}$$



Resultados de Medición

- Afirmación de conocimiento acerca de una o más cantidades atribuidas al sistema bajo medición
- Formuladas en clave de objetos científicos, conceptos abstractos y universales –e.g. masa, corriente, temperatura, duración, pobreza

Aspecto material
(interacción concreta, conocida)

Fenómenos
(ante los ojos)

≠

Observación
(codificada)

Datos

Subdeterminados
(componente interpretativo)

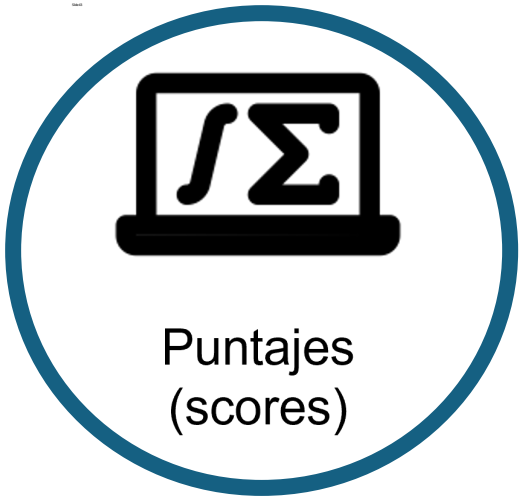
Fenómenos
(ante los ojos)

≠

Objetos científicos

≠

Estimadores



Aspecto epistémico
(representación abstracta)

Puntajes

≠

Objetos científicos

Indicaciones vs resultados

Indicaciones instrumentales (lecturas; propiedad del instrumento en su estado final)

- Volumen de la columna de mercurio en un termómetro
- Posición de una aguja en relación con el dial de un amperímetro
- Número de ciclos ("tics") generado por un reloj

Resultados de medición (estimado del valor de una cantidad bajo medición, con incertidumbre asociada)

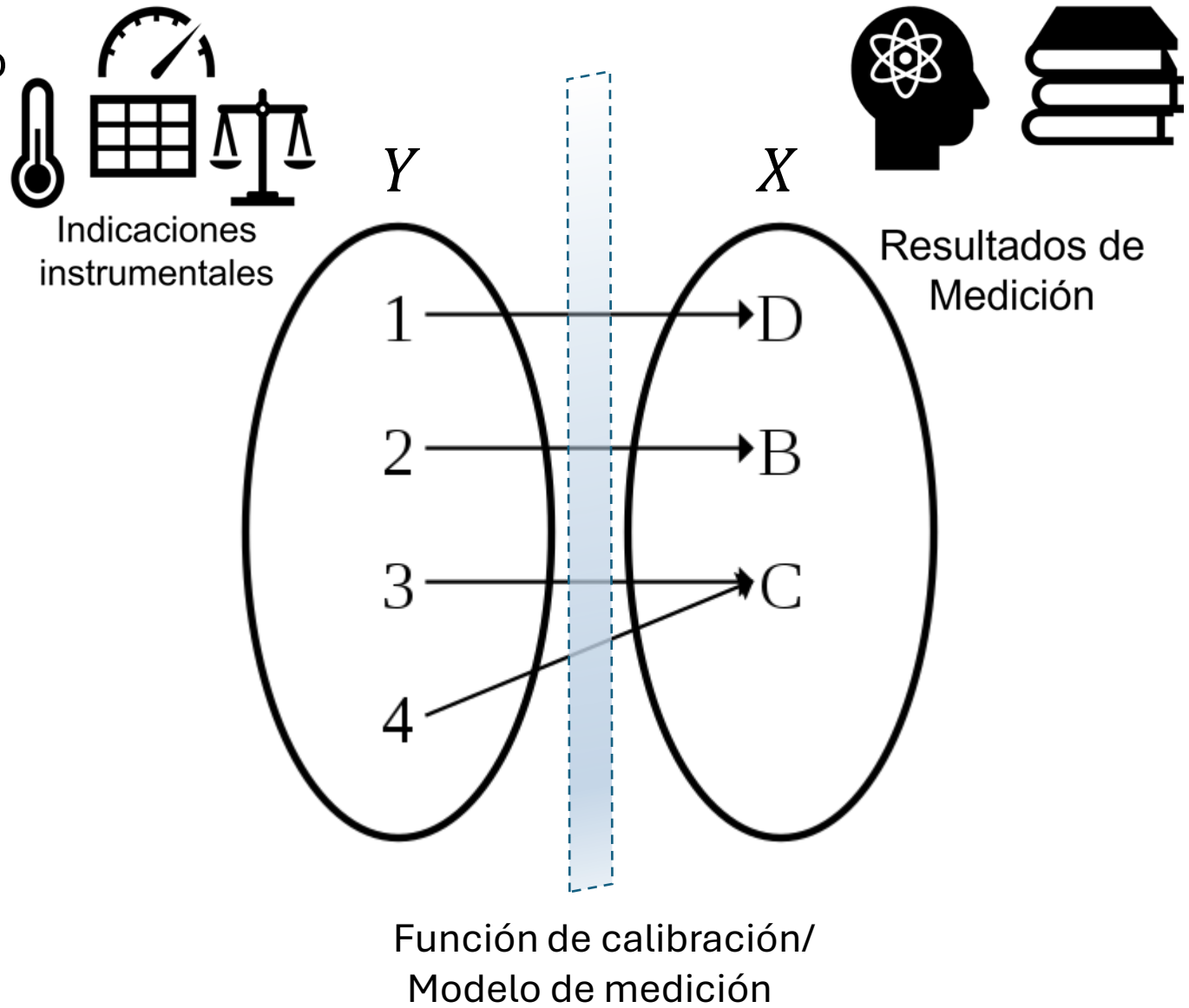
- Temperatura estimada con incertidumbre
- Corriente eléctrica estimada con incertidumbre
- Duración estimada con incertidumbre

Naturaleza inferencial de la medición

Los resultados de medición sólo pueden ser inferidos una vez que el instrumento ha sido subsumido bajo un modelo idealizado que les relaciona (son modelo-dependientes)

Medir es hacer inferencias

- Queremos medir la cantidad X (concepto abstracto y universal)
- La cantidad X no es directamente observable, así que inferimos de otra cantidad Y, que es "directamente observable" (indicaciones instrumentales).
- Para llevar a cabo esta inferencia necesitamos una regla de correspondencia que exprese X como función de Y como sigue $X = f(Y)$
- La forma de esta función f no puede ser descubierta ni estimada, debido a que involucra saber los valores tanto de Y como de X, y X es la variable desconocida que estamos tratando de medir.



Fenómenos
(ante los ojos)

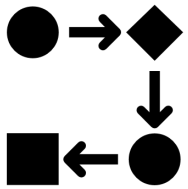
≠

Objetos científicos

Fenómenos
(ante los ojos)

≠

Observación
(codificada)



Modelo de
medición

Puntajes

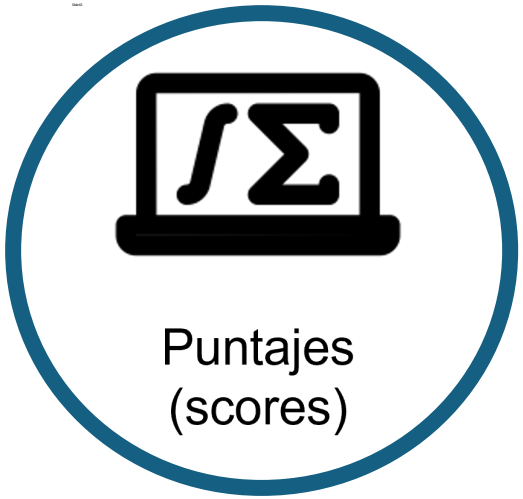
≠

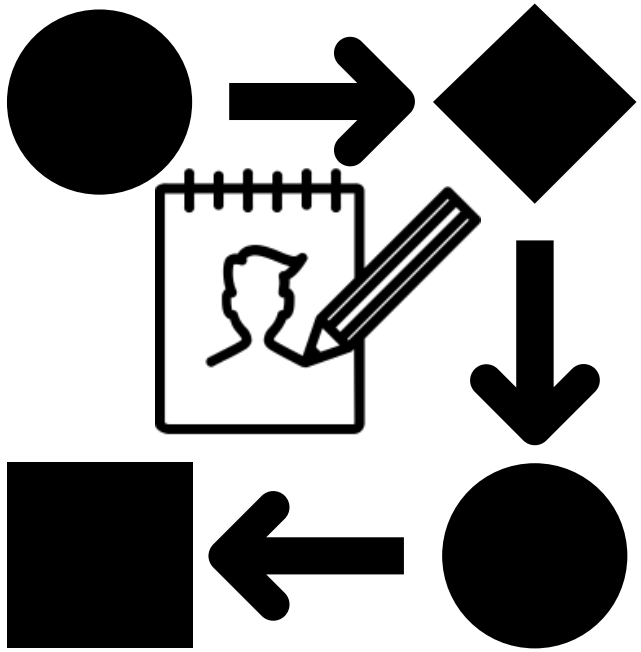
Objetos científicos

Datos

≠

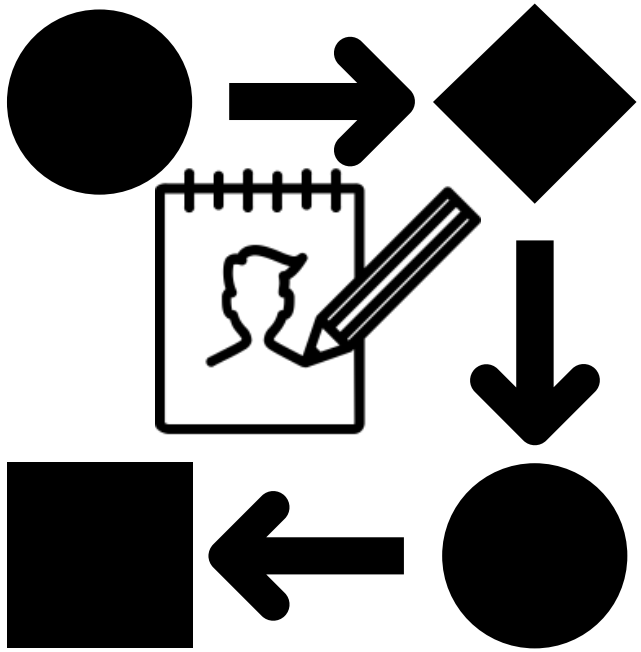
Estimadores





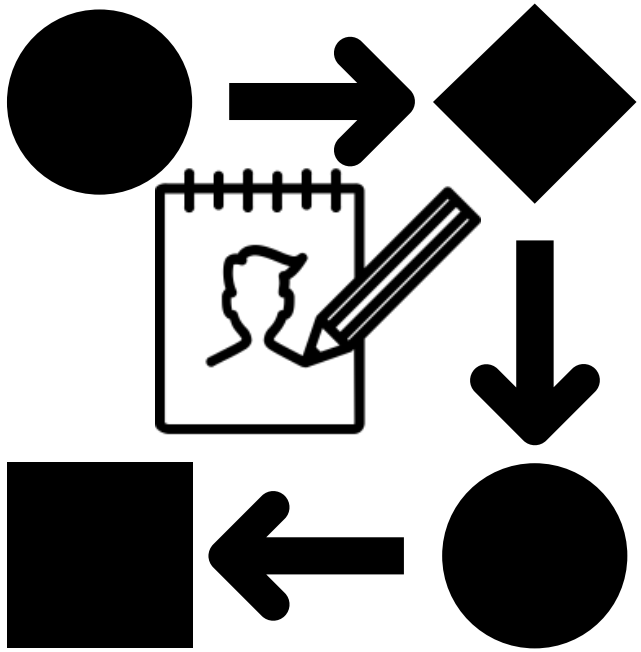
Modelo de medición

- Una representación abstracta y local construida a partir de supuestos simplificadores
- Mediador entre los niveles material y epistemológico
- Hipótesis teóricas sobre las relaciones que guardan los instrumentos con aquello que se quiere medir y con el ambiente ([DAG] sobre cómo fueron producidos los datos)
- Modelo teórico o estadístico del proceso de medición mismo



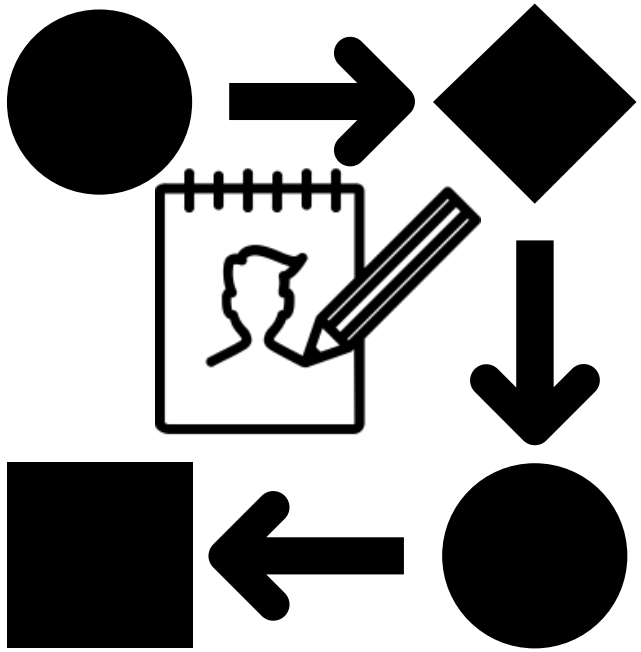
Modelo de medición

- Un papel reconocido ya por Duhem (1914), Kuhn (1961) y Suppes (1962)
  - “Si el experimento de física fuera la simple constatación de un hecho, sería absurdo introducir en él correcciones. Una vez que el observado hubiera mirado atenta, cuidadosa y minuciosamente, sería ridículo decirle: lo que ha visto no es lo que debería haber visto; permítame que haga unos cálculos que le enseñarán lo que debería haber constatado” Duhem (1914)



Modelo de medición

- Descripción transparente del sistema físico de transmisión de información (cómo son producidos los datos)
- Permite la rastreabilidad/trazabilidad de la generación de los resultados de la medición (a lo largo de cada eslabón de la cadena) en su relación con aquello que se quiere medir
- Establece relaciones **cuantitativas** entre aquello que se quiere medir y el resultado de su medición
- Generativos: genera instancias de datos (input-output de acuerdo con el proceso de medición idealizado)

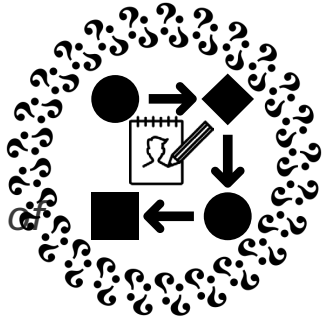


Modelo de medición

- Indispensable para hablar de error en la medición: la discrepancia entre el valor obtenido bajo el proceso de medición ideal (lo que debió haberse observado) y el obtenido con el proceso de medición que de hecho tuvo lugar (lo observado)
- Sólo bajo el modelo es posible evaluar la *interpretabilidad representacional* de los puntajes (su validez)
  - Coherencia de los supuestos con las teorías contextuales relevantes
  - Consistencia mutua de resultados con diferentes instrumentos, ambientes y

Confiabilidad y validez

Sólo bajo el modelo de medición



- *1.* Mental health screening of preschool children: Validity and reliability of ABLE. *American Journal of Orthopsychiatry*, *77*(3), 402–418, 402.

  – *ABLE (Attention, Behavior, Language, and Emotions), a new screening tool, was used to estimate the prevalence and the severity of concerns parents and teachers have about children's school adjustment and evaluate their need for services. Data obtained from parents and teachers of children randomly selected from public Pre-K classrooms in 6 states (N = 415) and from a mental health screening of rural and urban children (N = 5,577) support the validity and reliability of ABLE*. Barbarin, O. A. (2007).

- *2.* Hides, L., Dawe, S., Young, R. McD., & Kavanagh, D. J. (2006). The reliability and validity of the Severity of Dependence Scale for detecting cannabis dependence in psychosis. *Addiction*, *102*, 35–40.

  – *The SDS [Severity of Dependence Scale] is a brief, valid and reliable screen for cannabis dependence among people with psychosis*.

Trivializar la medición

Afirmaciones acerca de que

- Aquello que se quiere medir es observable
- La medición es un tipo de observación en sí misma que no necesita de modelaje

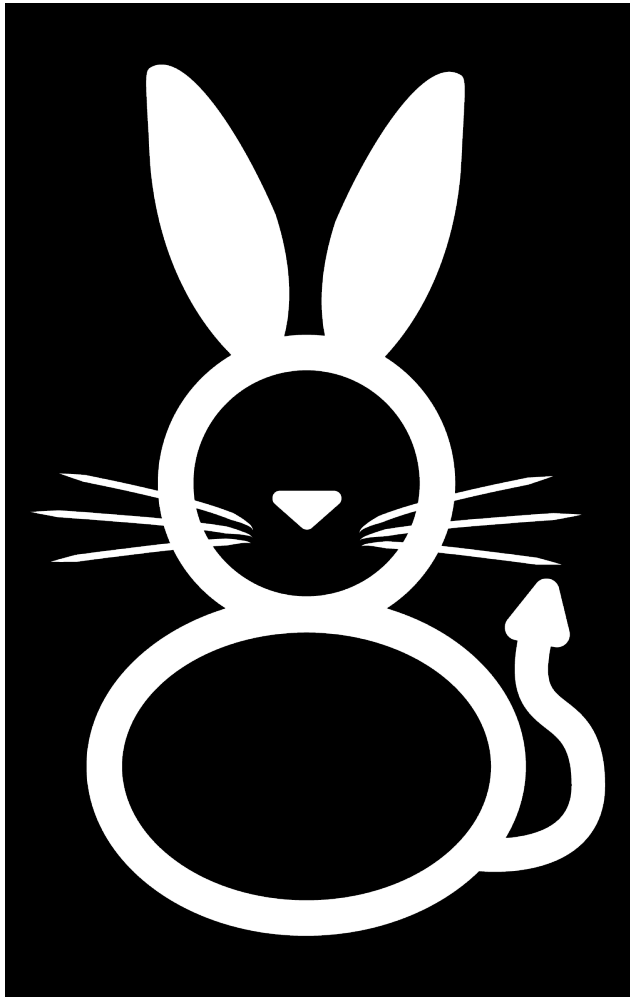
Reducen el significado de aquello que se quiere medir a un conjunto de datos y las operaciones involucradas

=

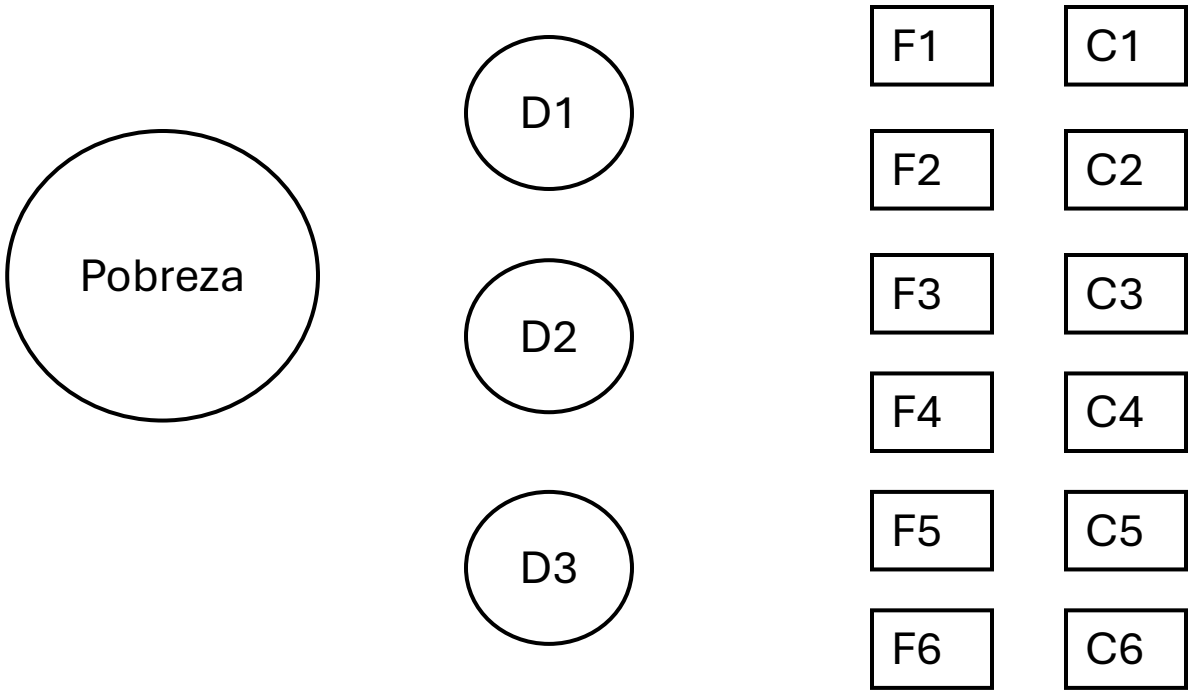
- Ausencia de error de medición
- Interpretabilidad representacional automática

Trivializar la medición

- Compromete la generalización científica (interpretaciones comunes de los resultados, independientemente de los procedimientos específicos)
- Comparaciones cuantitativas, geográficas y en el tiempo, pueden confundirse con aquello que queremos medir
- Diferencias espurias, enmascarar diferencias genuinas, conclusiones incorrectas



¿Hay modelo teórico?



¿Son finitas las Ds, Fs y las Cs?

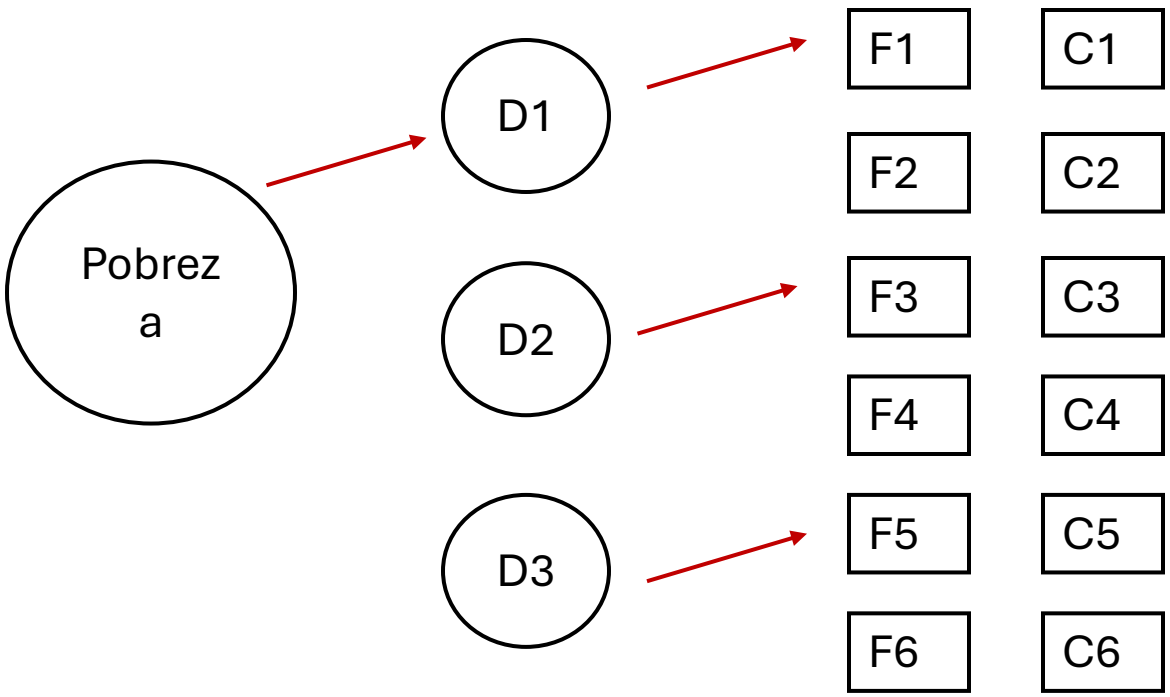
¿Son equivalentes las Fs y las Cs?

¿Cómo se relacionan las Fs y las Cs con las Ds?

¿Qué hay de las diferentes tasas de transformación?

¿F1 y F2 entre personas?

¿Hay modelo teórico?



¿Son finitas las Ds, Fs y las Cs?

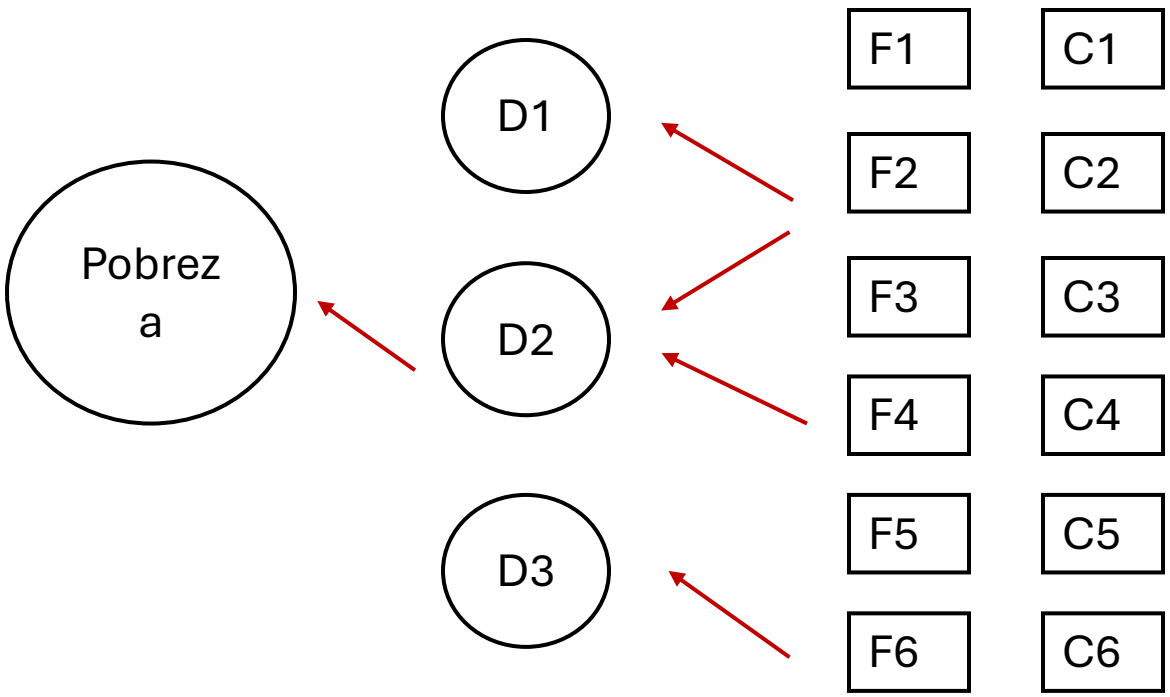
¿Son equivalentes las Fs y las Cs?

¿Cómo se relacionan las Fs y las Cs con las Ds?

¿Qué hay de las diferentes tasas de transformación?

¿F1 y F2 entre personas?

¿Hay modelo teórico?



¿Son finitas las Ds, Fs y las Cs?

¿Son equivalentes las Fs y las Cs?

¿Cómo se relacionan las Fs y las Cs con las Ds?

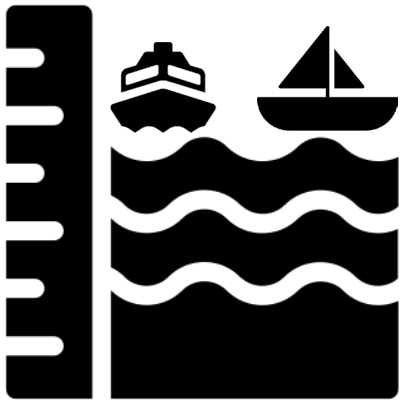
¿Qué hay de las diferentes tasas de transformación?

¿F1 y F2 entre personas?

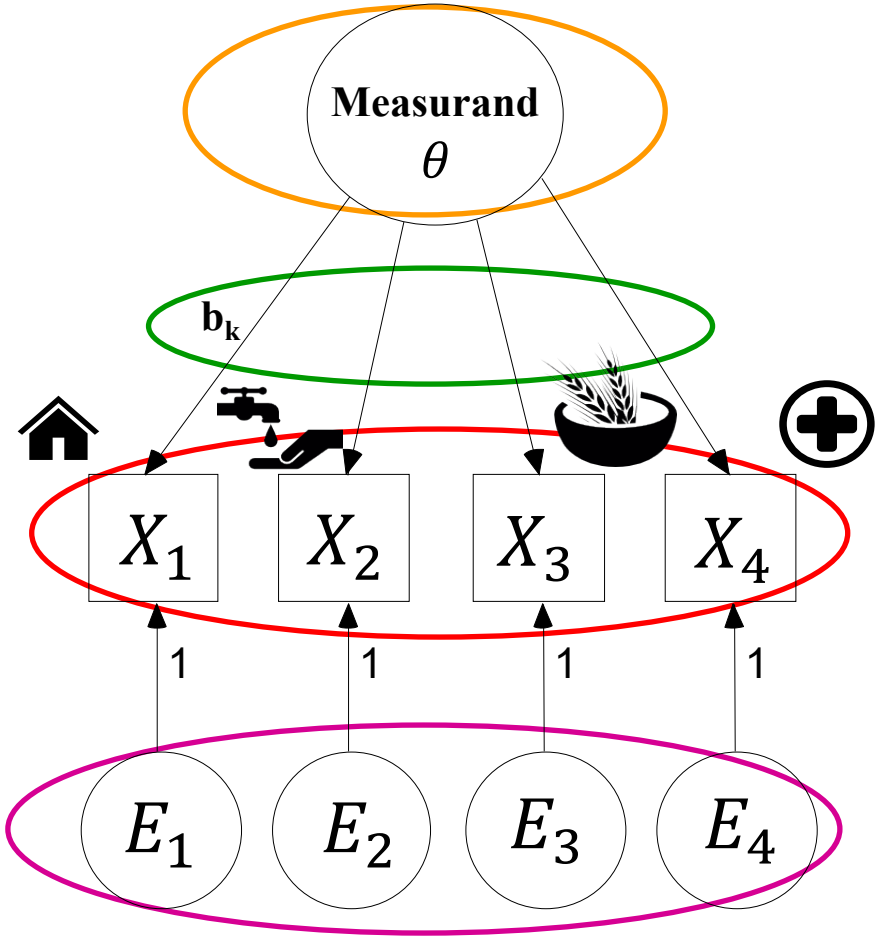
Mismos problemas que en NBI

SEM y medición

- Aquello que se quiere medir existe a un nivel conceptual de mayor profundidad que los indicadores
- Los indicadores son su consecuencia



$$X_{ik} = [\ a_k + b_k\ (\text{Measurand}_i)\ ] + E_{ik}$$



The current members of the Council of the Society are:

President

Luca Mari

Università Carlo Cattaneo – LIUC
Term expires in 2028

2026 Conference Host (ex officio)

Leah McClimans

University of South Carolina
Term expires in 2028

Council Member

Nadine de Courtenay

Université Paris Diderot
Term expires in 2026

Council Member

Mark Wilson

University of California, Berkeley
Term expires in 2028

Secretary

Eran Tal

McGill University
Term expires in 2026

Council Member

Jo Wolff

The University of Edinburgh
Term expires in 2026

Council Member

Miguel Ohnesorge

University of Cambridge
Term expires in 2026

Web Coordinator

David Torres Irribarra

Pontificia Universidad Católica de Chile
Term expires in 2028

2026 – Second Biennial Conference

June 22-25 2026

University of Edinburgh

We are excited to announce that the Second Biennial Conference of the Society for the Study of Measurement will be held at the University of Edinburgh June 22-25, 2026. The main conference will take place June 23-25, with a pre-conference day of workshops held on June 22.We are delighted to announce that **Professor Jana Uher** (Greenwich) will be our keynote speaker and that **Professor Luca Mari** (Università Carlo Cattaneo – LIUC) will be giving the society's inaugural presidential address.

Call for Proposals

We invite proposals on any topics in the theory, history, philosophy, and application of measurement. We especially welcome proposals connected to this year's theme *Ground Truth and Validity*. Proposals may take any of the following forms:

- **Individual papers for a 20-minute presentation (+10 minutes Q&A)**
- **Symposium of 2-4 papers for a 90-minute session**
- **Posters**
- **Workshops for 2-4 hours on June 22nd**

Please submit proposals for individual contributions (~500 words) or symposia/workshops (~1,200 words), specifying the chosen format (a)-(d), and following the instructions of the website. Additional information and the forms to submit your proposal can be found at the dedicated submission site.

Full submission link: https://app.oxfordabstracts.com/stages/80364/submitter

Please feel free to share this announcement with colleagues.

We look forward to seeing you in Edinburgh in June 2026!

Conclusiones y discusión

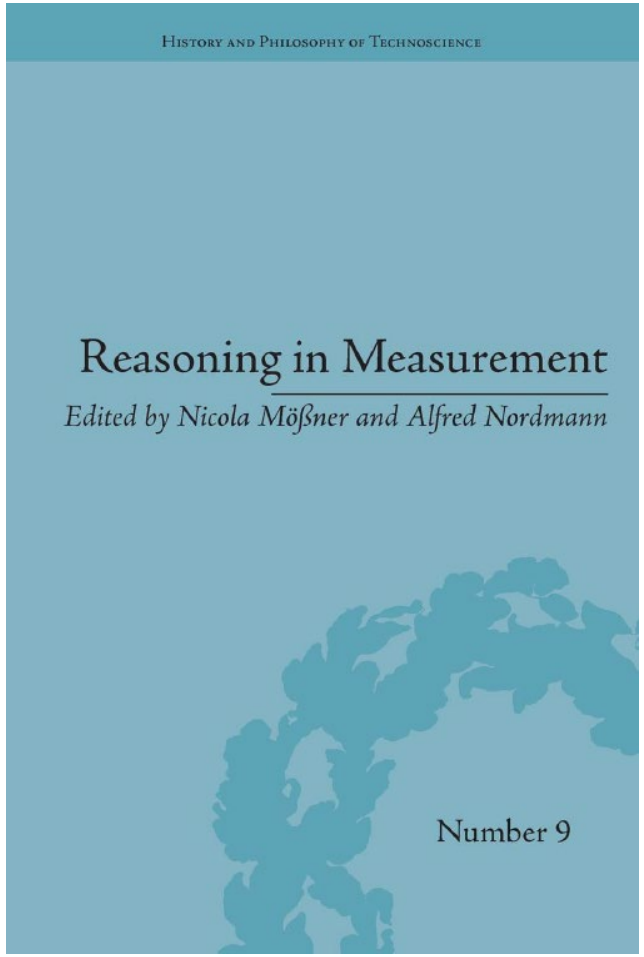
Aproximación basada en modelos:

- Trasciende los ejercicios de crítica horizontal entre distintas propuestas de medición.
- Coloca a las discusiones semánticas (robustez v validez) o técnicas (método A v método B) en el plano de las hipótesis y producción de evidencia
- Ofrece un marco para establecer el tipo de preguntas que debería contestar cualquier ejercicio de medición
- Establece como condición necesaria explicitar los términos bajo los que se define y calcula el error de medición

De esta sesión

Epistemología de la medición

- Tal, E. (2017) "A Model-Based Epistemology of Measurement", in Mößner, N. and Nordmann, A. (eds.), *Reasoning in Measurement*, London and New York: Routledge, pp. 233-253.



De esta sesión

Problemas del Desarrollo. Revista Latinoamericana de Economía, vol. 55, núm. 217, abril-junio 2024.

WHAT WE TALK ABOUT WHEN WE TALK ABOUT MEASUREMENT IN POVERTY RESEARCH

Curtis Huffman Espinosa and Héctor E. Nájera Catalán[a]

Fecha de recepción: 13 de junio de 2023. Fecha de aceptación: 21 de diciembre de 2023.

https://doi.org/10.22201/iiec.20078951e.2024.217.70081

Abstract. It is hardly contentious to assert that measuring concepts is an important aspect of scientific and practical research. However, there seems to be some degree of confusion nowadays regarding what is meant by measurement in poverty research. It is obvious from these exchanges that different notions of central terms in the debate are being held (reliability, validity, measurement error, measurement model) to the detriment of common understanding. To move the literature forward, this paper falls back on the epistemology of measurement to bridge the apparent conceptual gap in the debate. This article invites more discussion and constructive exchange of views regarding the meaning of measurement in poverty research and how to assess the relative success of different efforts.

Key Words: welfare; well-being and poverty; measurement and analysis of poverty; economic methodology (general).