

# Introducción a los modelos mixtos

**Eduardo Garcia-Garzon**

IA Specialist



shakers

# Introducción

1. Qué son los modelos mixtos.
2. Formulación matemática.
3. Ventajas vs modelos “tradicionales”.
4. Limitaciones.
5. Tipos de modelos & ejemplos.

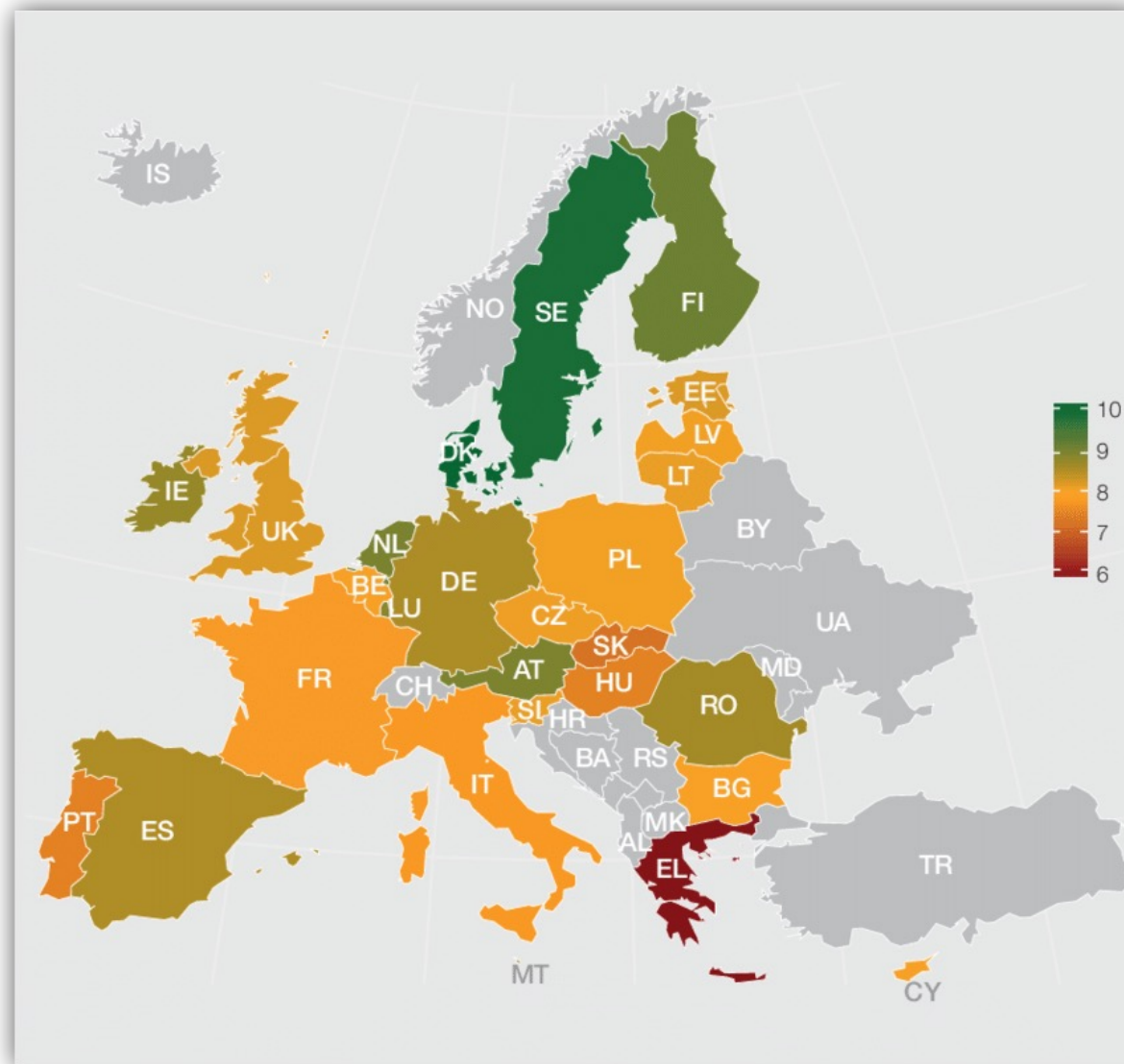
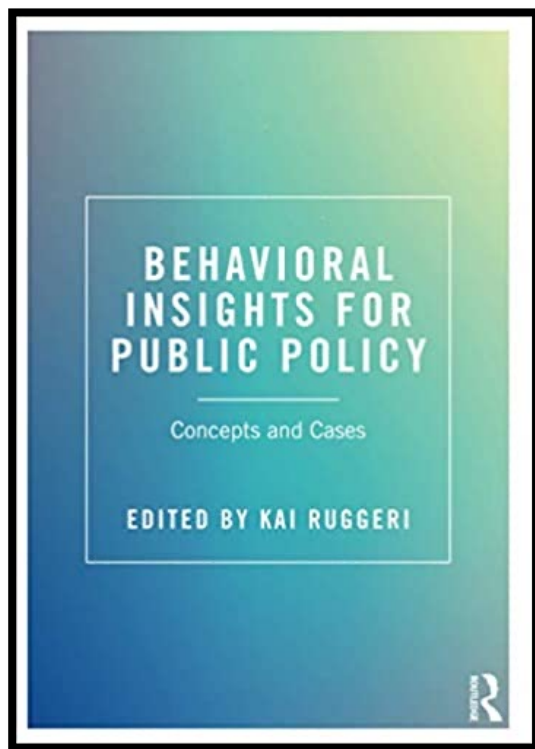


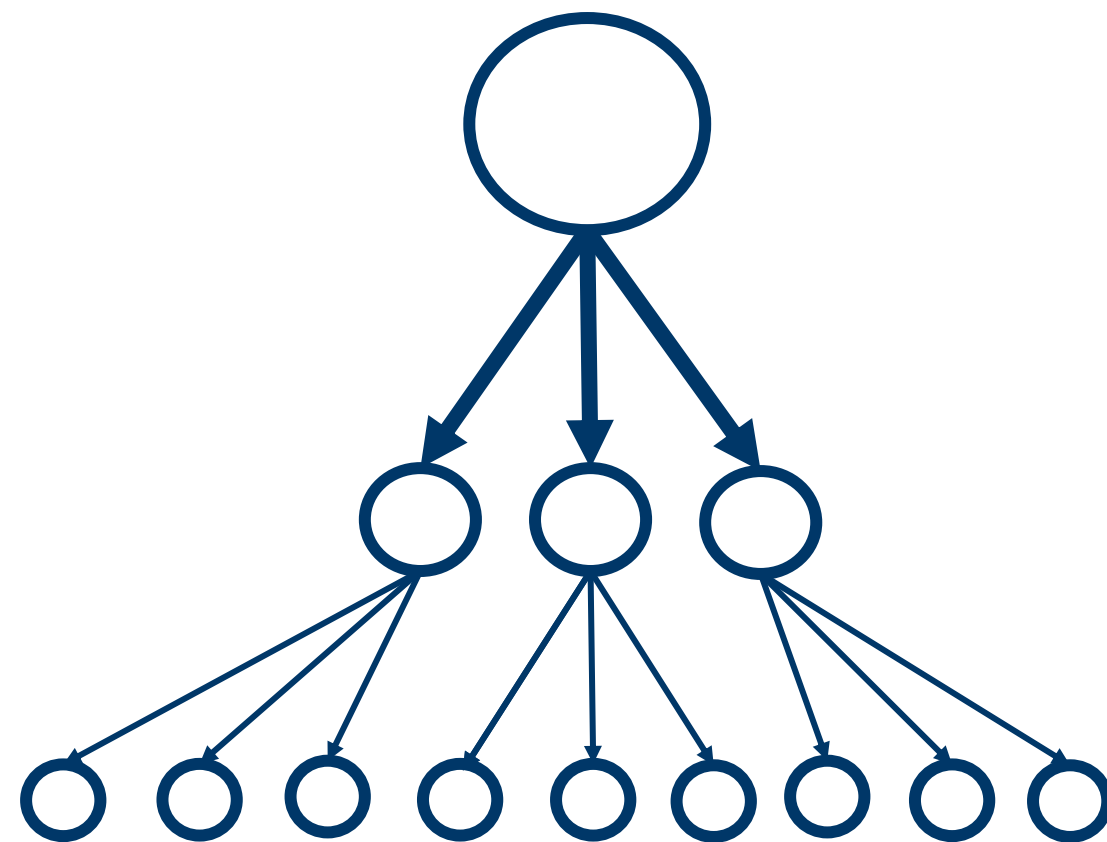
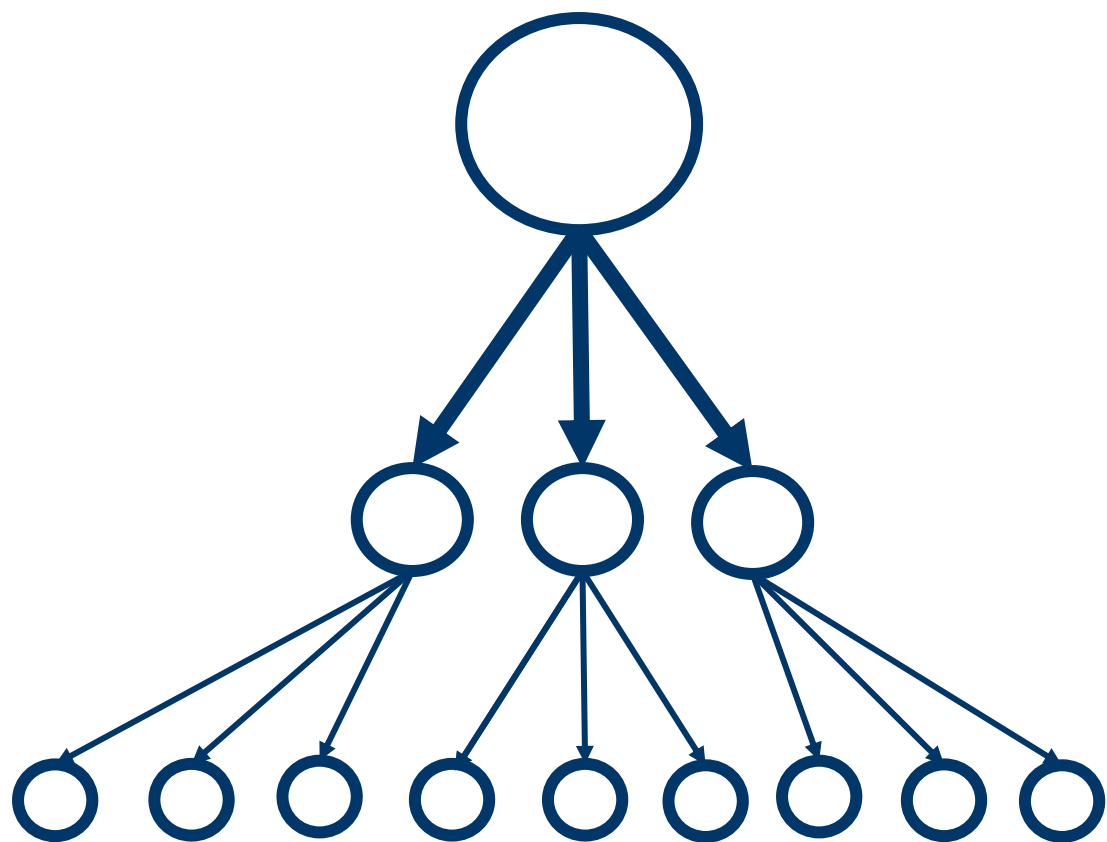
## Key messages

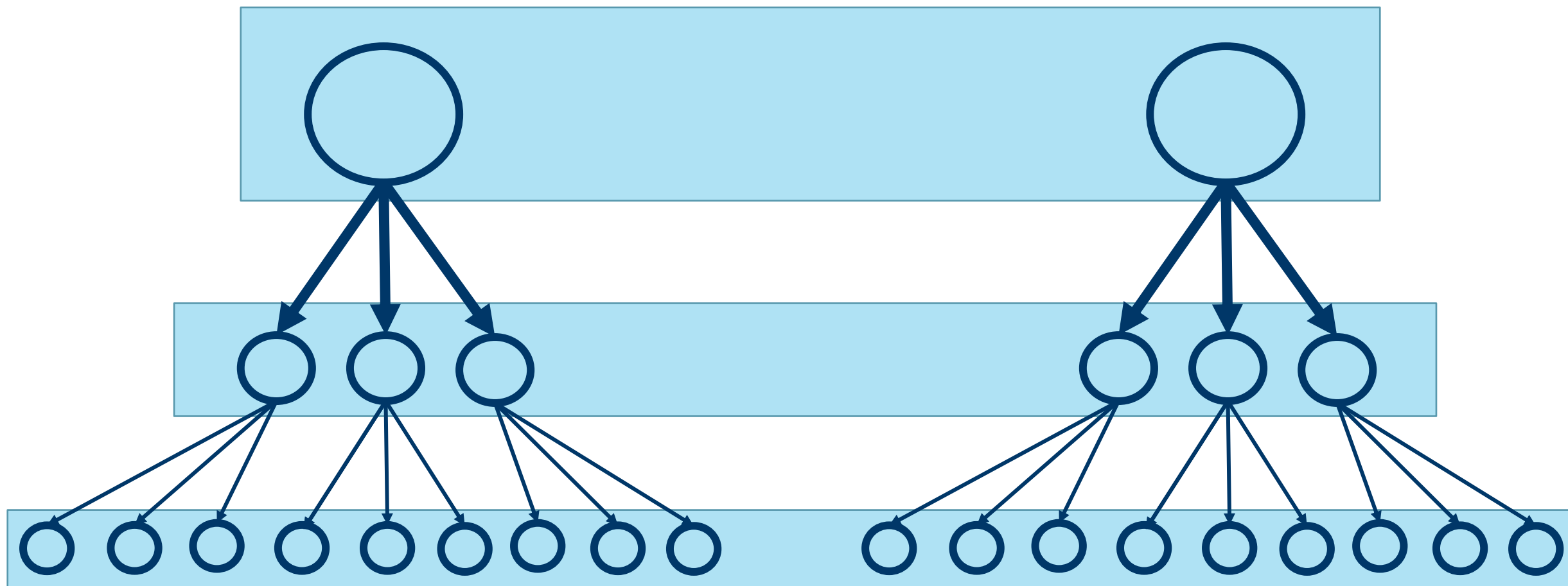
1. Qué es un efecto aleatorio
2. Qué beneficios tienen los MLM.
3. Qué tipos de MLM existen.

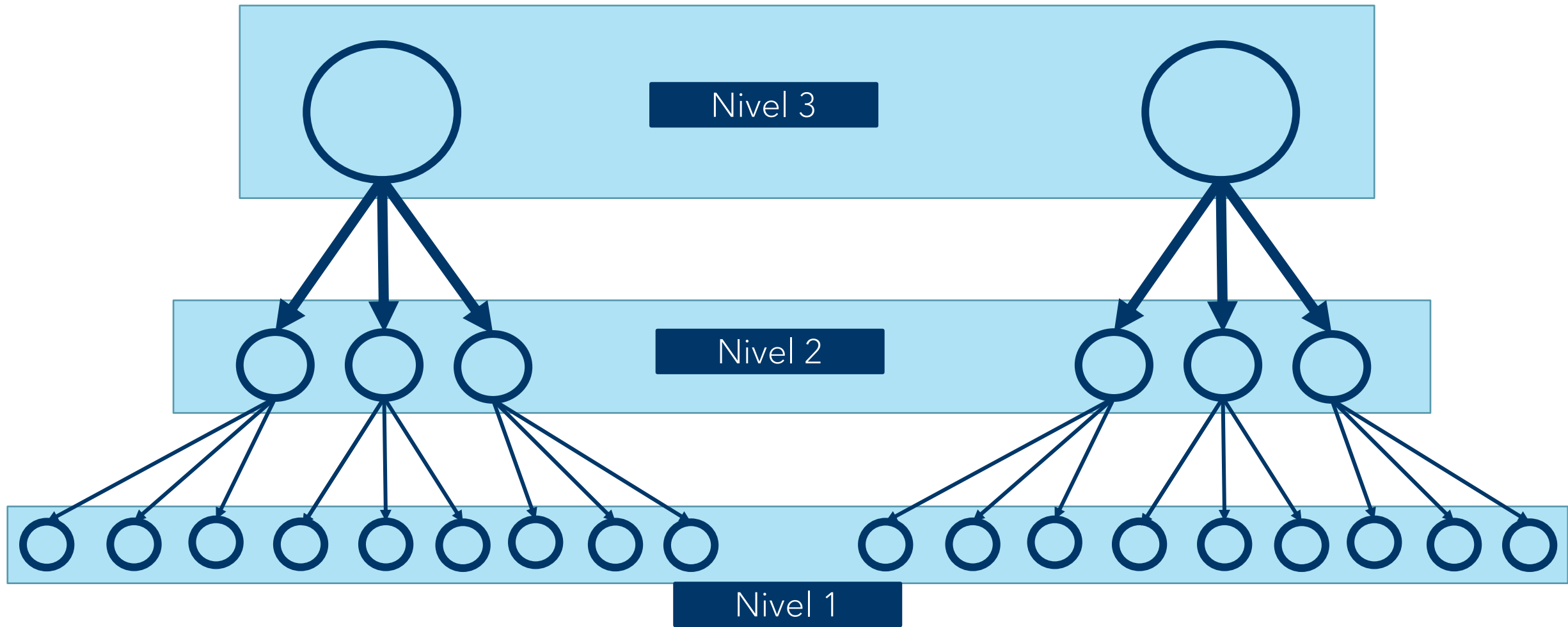


Modelos mixtos,  
bienvenidos

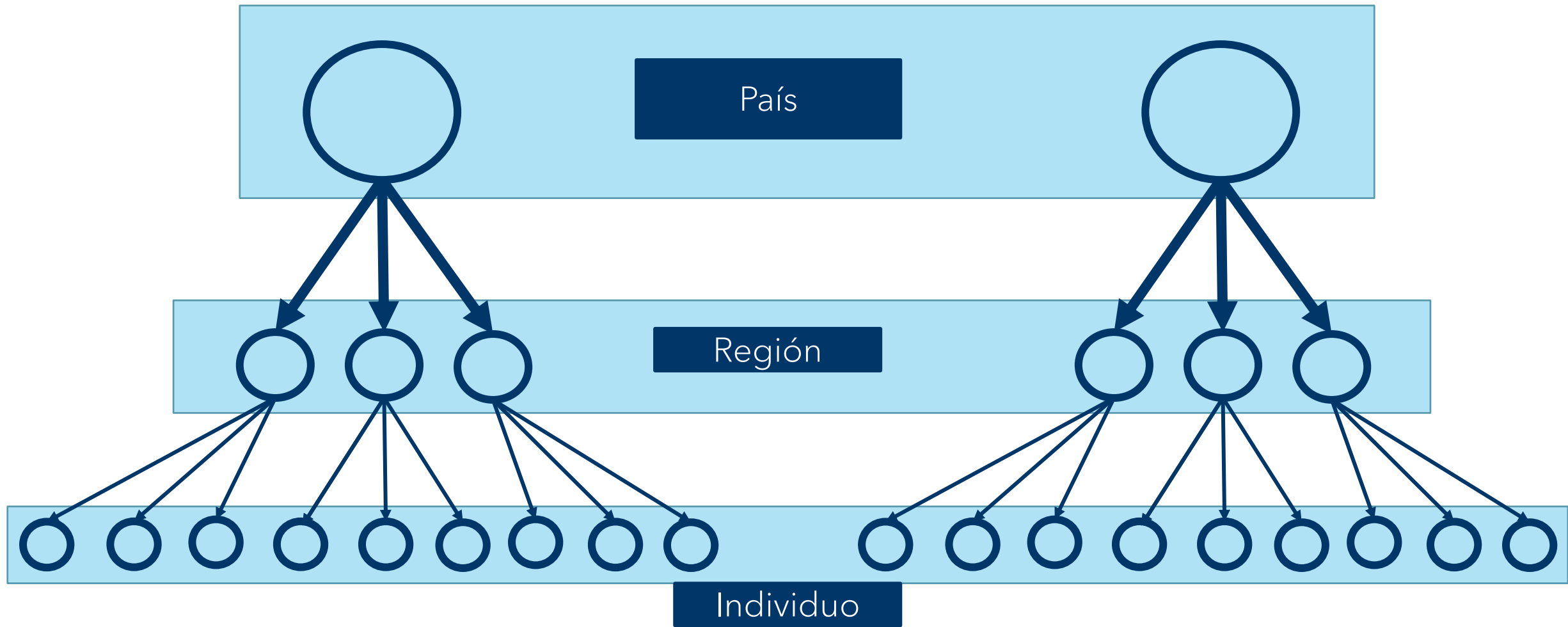


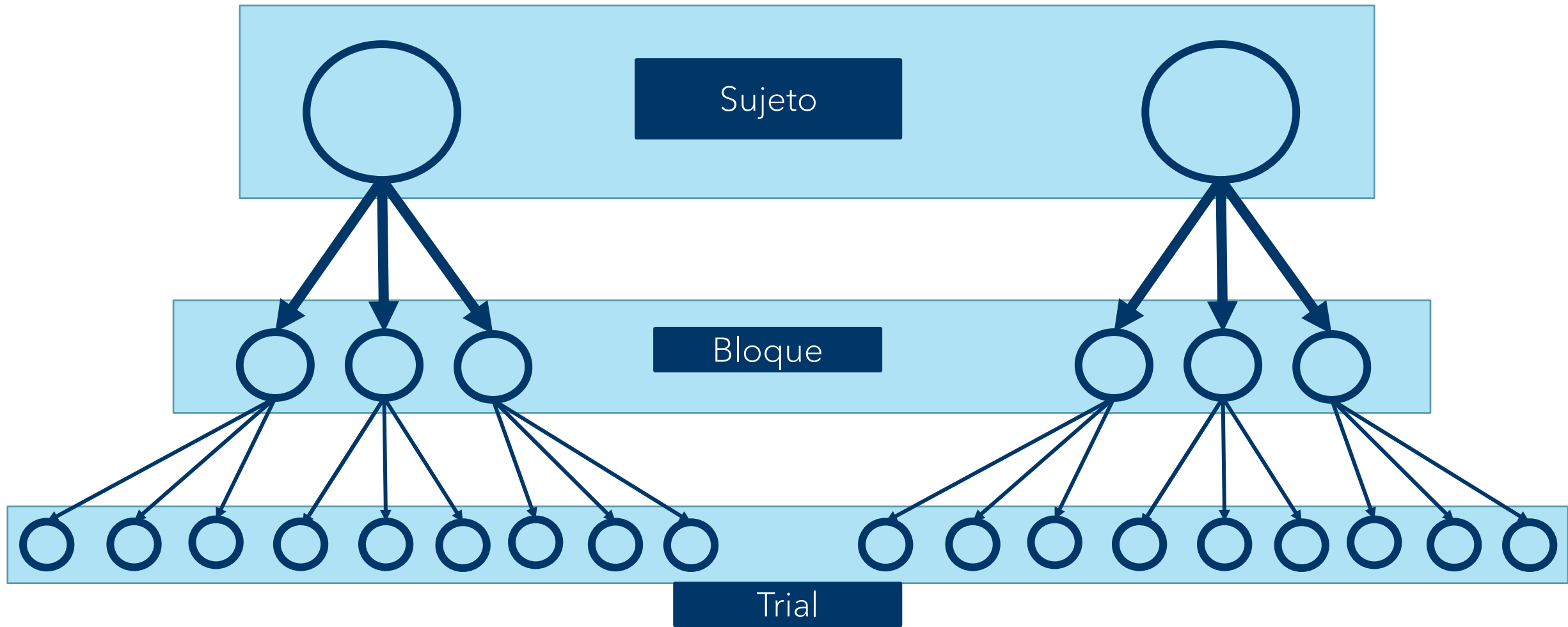


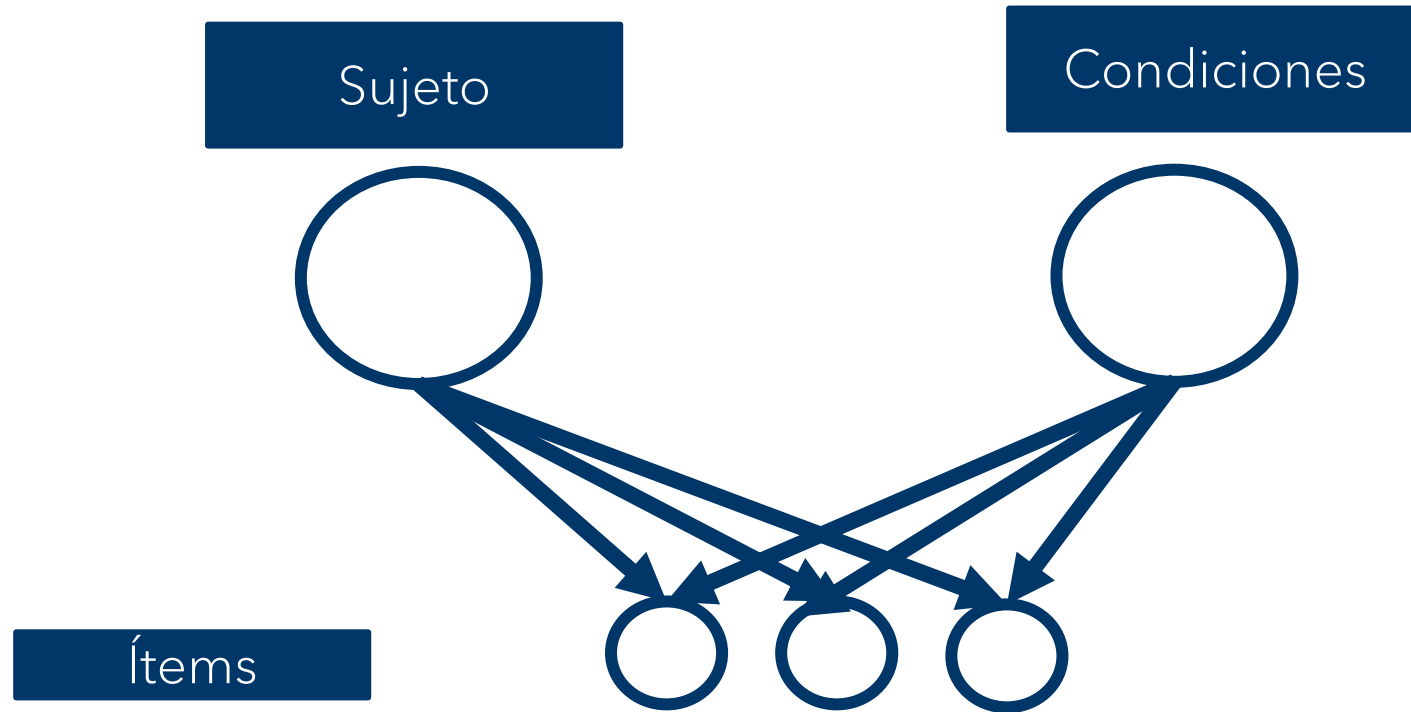




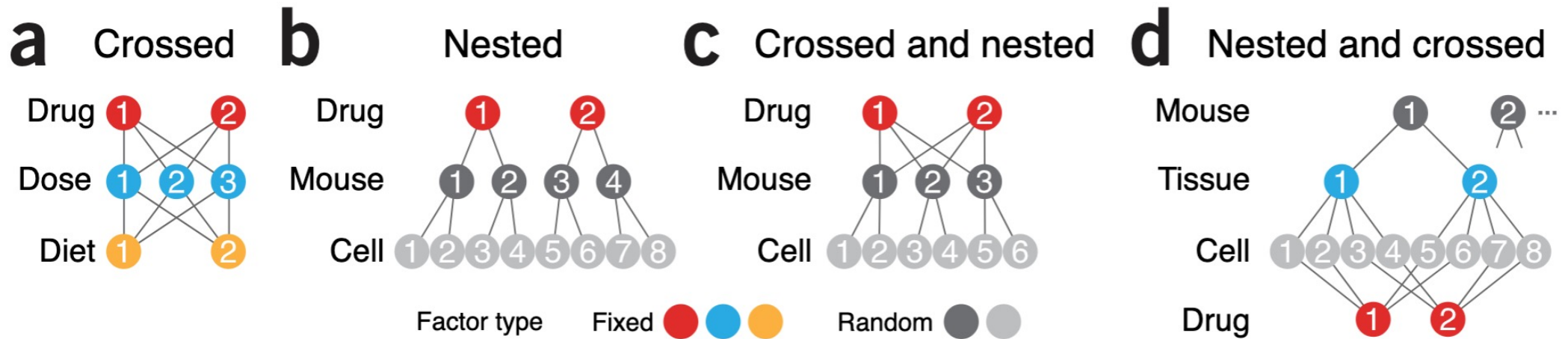








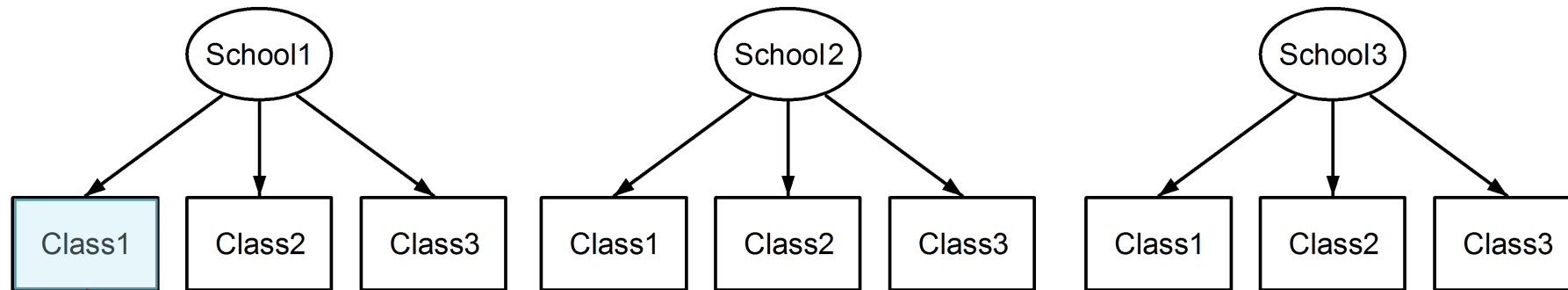
# La estructura de los datos define el tipo de modelo



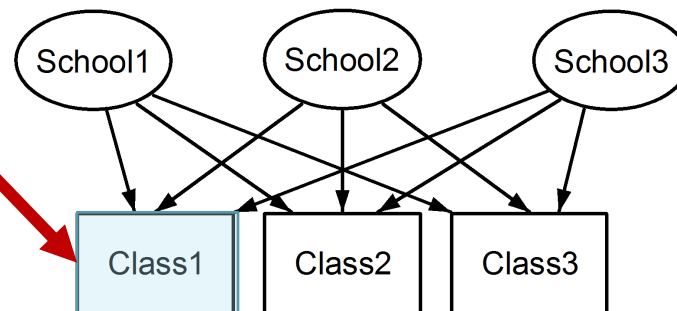
<https://www.nature.com/articles/nmeth.3137.epdf>



## Modelo jerárquico o anidado



## Modelo cruzado



<https://stats.stackexchange.com/questions/228800/crossed-vs-nested-random-effects-how-do-they-differ-and-how-are-they-specified>

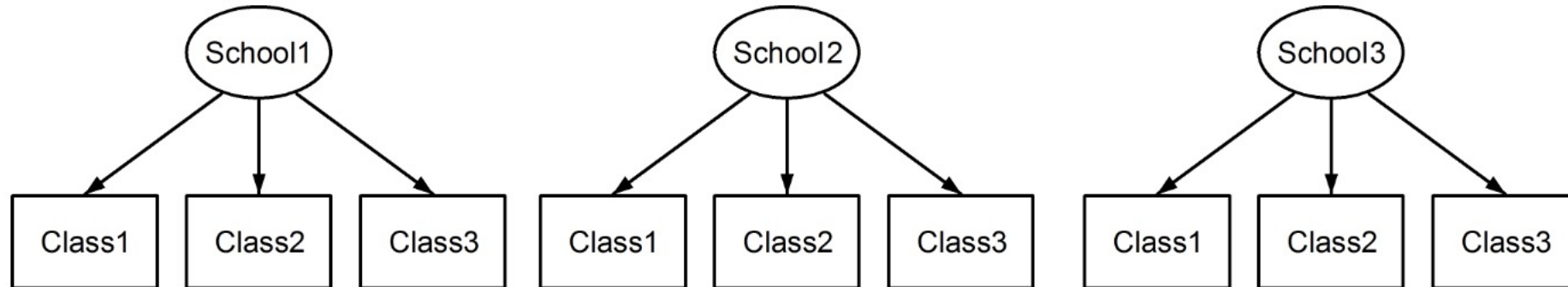


## ¿Ante qué tipo de efectos estaríamos?

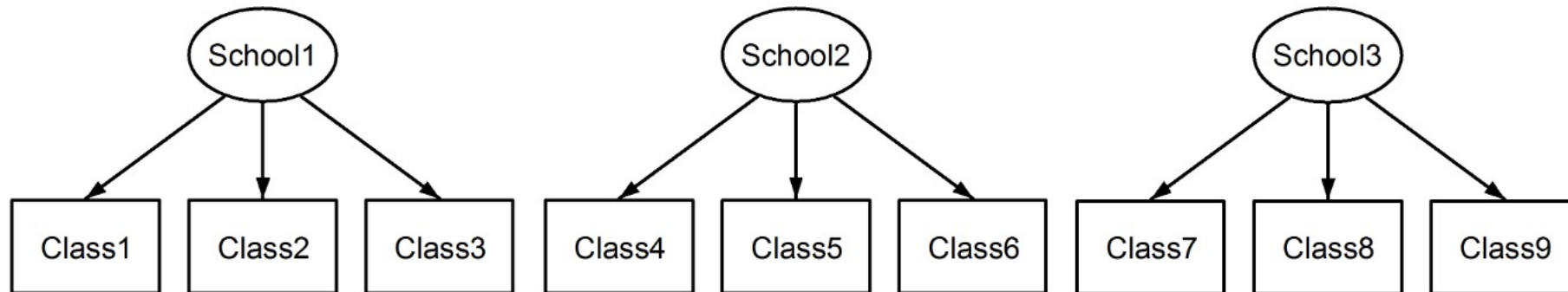
- Respuestas a una encuesta realizada en 5 países europeos sobre cambio climático.
- Respuestas a una encuesta realizada **anualmente (2019-2022)** en 5 países europeos sobre cambio climático.
- Respuestas a un test de ansiedad de individuos divididos al azar en experimental o un grupo control.
- Respuestas a un test de ansiedad de individuos bajo dos condiciones (experimental y control)



With this naming convention, nested and fixed effects will be **different**:



With this naming convention, nested and fixed effects will be **equivalent**:



<https://errickson.net/stats-notes/vizrandomeffects.html>

## No liarse.

- Si tenéis un modelo jerárquico/anidado -> generar un id único a cada unidad del nivel inferior para convertirlo en cruzado.
- Si el diseño es cruzado → no es necesario hacer nada.
- Aunque facilita la sintaxis posterior, no es común encontrar datasets tratados de este modo.





## En otras palabras

Estas estructuras reflejan que las unidades de información **no son independientes** a través de los niveles.

*En castellano:* es esperable que los individuos de la misma región se parezcan a las de personas de su misma zona, o que las respuestas del mismo individuo sean similares entre sí.

Si no tenemos en cuenta estas dependencias, la estimación de los coeficientes y los s.e. **será incorrecta**.



## Clave

¿Cómo traducimos esta intuición a un modelo estadístico?  
Utilizando **dos tipos** de coeficientes de regresión:

- 1) Coeficientes **fijos**.
- 2) Coeficientes **aleatorios**.

Un **modelo mixto** es un modelo de regression que presenta estos dos tipos de coeficientes (independientemente de su estructura).



## Modelos mixtos o...

- Modelos de factores aleatorios
- Modelos de coeficientes aleatorios
- Modelos jerárquicos\*
- Modelos multinivel\*
- Modelos anidados\*
- Modelos lineales mixtos\*
- Modelos de curvas de crecimiento.

Múltiples nombres, modelos parecidos (nomenclatura).

## ¿Por qué?

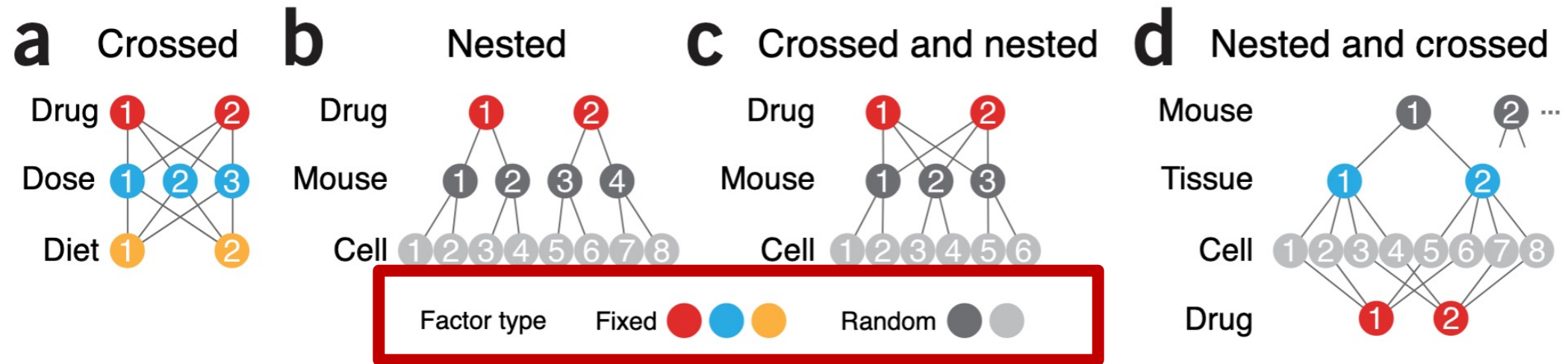
Los **modelos mixtos** gozan de una popularidad extrema en ciencias del comportamiento.

Hoy en día son una herramienta estadística **estándar** en un muchas comunidades científicas.

¿Por qué es necesario su uso? ¿Cuáles son sus ventajas? ¿Hay inconvenientes? ¿Con esto me dejarán en paz los revisores?



La estructura de los datos  
define el tipo de modelo



<https://www.nature.com/articles/nmeth.3137.epdf>



# Mi primer modelo mixto

## Que sería de un taller de metodología sin ecuaciones....

En el curso de psicometría, los **alumnos** acuden a tutorías opcionales de repaso. ¿Son efectivas estas tutorías para mejorar las notas en la asignatura?

Intentamos predecir el efecto de acudir a tutoría sobre la nota final teniendo en cuenta las diferencias en las notas promedio de los alumnos.



## Que sería de un taller de metodología sin ecuaciones....

Intentamos predecir el efecto de acudir a tutoría sobre la nota final teniendo en cuenta las diferencias en las notas promedio de los alumnos.

$$Nota = b_{intercepto} + b_{tutoría} * tutoría + \varepsilon$$

Asumiendo que...

$$\varepsilon \sim N(0, \sigma)$$

$$\mu = b_{intercepto} + b_{tutoría} * tutoría$$
$$Nota \sim N(\mu, \sigma)$$



## Que sería de un taller de metodología sin ecuaciones....

Añadimos un efecto específico por estudiante

$$Nota = (b_{intercepto} + eff_{estudiante}) + b_{tutoría} * tutoría + \varepsilon$$

$$\varepsilon \sim N(0, \sigma)$$

$$eff_{estudiante} \sim N(b_{intercepto}, \tau)$$





## Que sería de un taller de metodología sin ecuaciones....

Añadimos un efecto específico por estudiante

$$Nota = b_{intercepto\_estudiante} + b_{tutoría} * tutoría + \varepsilon$$

Asumiendo que...

$$b_{intercepto\_estudiante} \sim N(b_{intercepto}, \tau)$$

$$\varepsilon \sim N(0, \sigma)$$



## Que sería de un taller de metodología sin ecuaciones....

Tradicionalmente, condensamos la notación a estas dos ecuaciones principales..

- $Nota = b_{intercepto\_estudiante} + b_{tutoría} * tutoría + \varepsilon$
- $b_{intercepto\_estudiante} \sim N(b_{intercepto}, \tau)$
- $\varepsilon \sim N(0, \sigma)$



## Que sería de un taller de metodología sin ecuaciones....

Tradicionalmente, condensamos la notación a estas dos ecuaciones principales..

$$\text{Nota} = b_{\text{intercepto\_estudiante}} + b_{\text{tutoría}} * \text{tutoría} + \varepsilon$$
$$b_{\text{intercepto\_estudiante}} \sim N(b_{\text{intercepto}}, \tau)$$



## Que sería de un taller de metodología sin ecuaciones....

Tradicionalmente, condensamos la notación a estas dos ecuaciones principales..

$$\begin{aligned} \text{Nota} &= b_{\text{intercepto\_estudiante}} + b_{\text{tutoría}} * \text{tutoría} + \varepsilon \\ b_{\text{intercepto\_estudiante}} &\sim N(b_{\text{intercepto}}, \tau) \end{aligned}$$

[http://mfviz.com/hierarchical-models/?fbclid=IwAR074-goLDIzwMwWmq86qINQRGIEclsm\\_SAfXin4JulC\\_RSGZU6MJIDBXTc](http://mfviz.com/hierarchical-models/?fbclid=IwAR074-goLDIzwMwWmq86qINQRGIEclsm_SAfXin4JulC_RSGZU6MJIDBXTc)



Que sería de un taller de metodología sin ecuaciones....

$$y = \beta_0 + \sum \beta_i x_i + \gamma + \varepsilon$$

Response variable

Global intercept

Fixed effect parameters

Fixed effect variables

Random effect variance

Residual variance

Silk et al. (2020), PeerJ, DOI 10.7717/peerj.9522



## Modelo de interceptos aleatorio (random intercept)

- El modelo previo fija que los interceptos de cada alumno son extraídos de una distribución → **modelo de interceptos aleatorios.**
- Es el modelo mixto más sencillo de los que vamos a construir.
- El resto de modelos mixtos se basan en definir coeficientes fijos como distribuciones (i.e., pendientes) y modelar la matriz de var-covar entre estos efectos.

[http://168.176.137.33:3838/fhernanb/lmm\\_b0b1/](http://168.176.137.33:3838/fhernanb/lmm_b0b1/)



¿Para qué necesito modelar un efecto aleatorio?

## Goodbye, ANOVA

- Todo ANOVA puede parametrizarse como un modelo de regresión lineal (pero lo contrario no es cierto).
- Los modelos mixtos son más generales, aportan más información, y son el método de-facto.
- Ej. Un modelo de intercepto aleatorio equivalente a un ANOVA-MR\*.
- ¡Ojo! Existen ANOVAs que incorporan efectos mixtos, pero su uso suele ser complejo y menos flexible.





## ANOVA vs Modelos Mixtos I

- ANOVA solo permite predictores continuos como ANCOVAS.
- Menor potencia estadística para efectos aleatorios.
- Su generalización es compleja (e.g., binarios, Poisson, etc.).
- Difícil incorporar efectos/var.dependientes no-lineales.

## ANOVA vs Modelos Mixtos II

- Problemas con datos perdidos.
- Problemas con clases inbalanceadas.
- Problemas con diseños complejos (crossed-effects o varios efectos aleatorios).
- Problemas con covariables específicas a efectos aleatorios



## ¿Merece la pena abandonar el ANOVA?

- En condiciones balanceadas (efectos/muestras), con estructuras jerárquicas, y variables dependientes continuas.... obtendremos resultados similares.
- La diferencia va a depender de la influencia de los factores aleatorios → reflejada en los errores típicos y los valores p.
- **No existe la necesidad de jugársela.**



## La decisión real

- El marco de análisis va a ser el modelo lineal general, no el ANOVA.
- ¿Decisión? El tipo de modelo, el tipo de efectos, supuestos...
- **La discusión del ANOVA vs modelos mixtos está superada en una mayoría de campos.**
- Asegurar que la técnica utilizada es adecuada.



# Primer ejemplo

## Formato de los datos

- En un modelo mixto, los datos tienen que estar en formato "wide" o "amplio".
- Esquema mental: cada fila es un dato único.
- **Tantas filas como mediciones del factor del nivel más bajo de nuestro diseño experimental.**
- Recomendación: siempre utilizar el formato largo.



## ANOVA-MR

Efectos Dentro de los Sujetos

	Suma de Cuadrados	gl	Media Cuadrática	F	p
MR Factor 1	9.72	6	1.62	1.53	0.167
Residual	630.74	594	1.06		

*Nota.* Suma de Cuadrados Tipo 3

[3]

## Modelo mixto

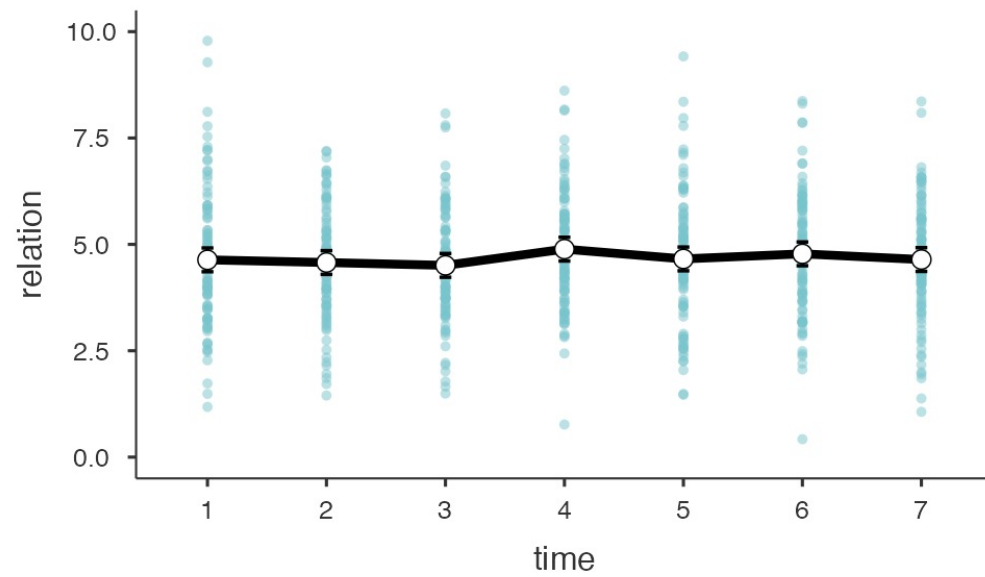
Fixed Effect Omnibus tests

	F	Num df	Den df	p
time	1.53	6	594	0.167

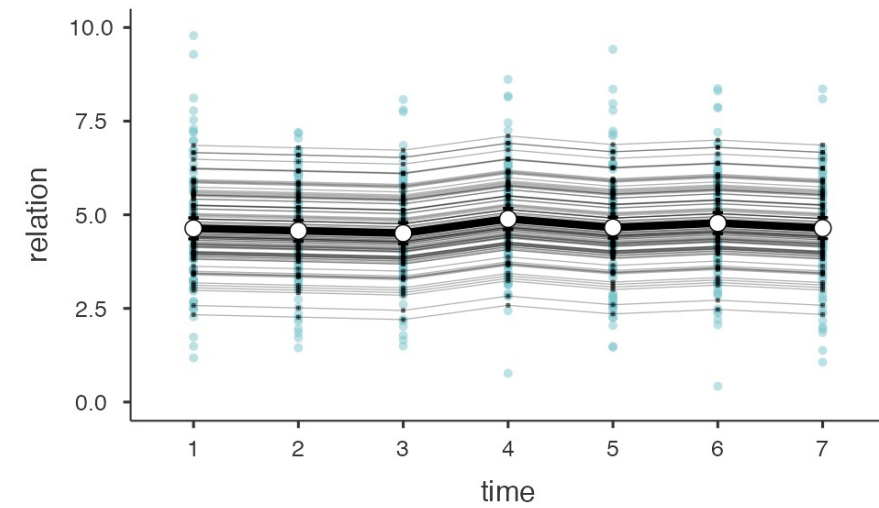
*Nota.* Satterthwaite method for degrees of freedom



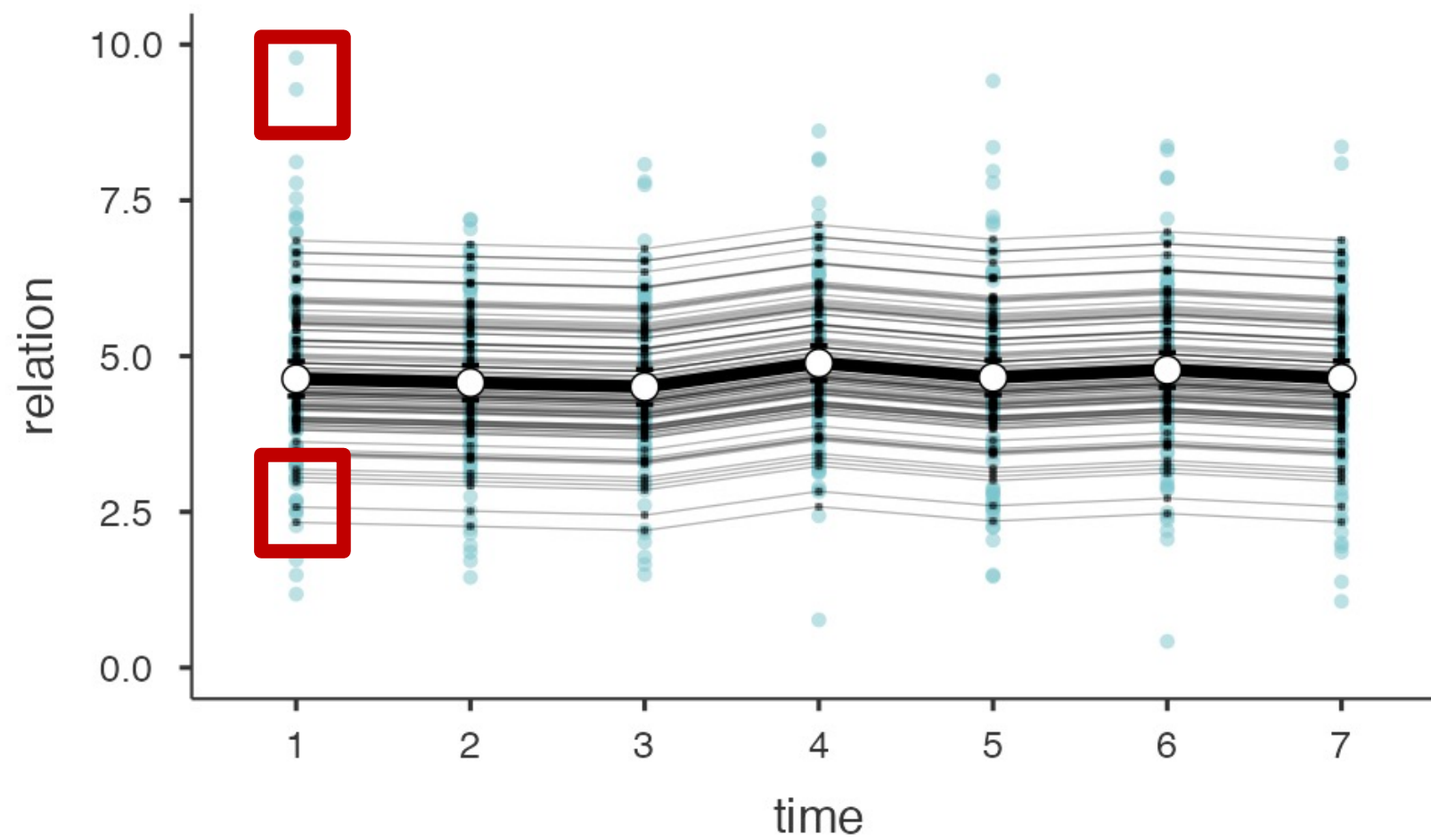
## Effects Plots



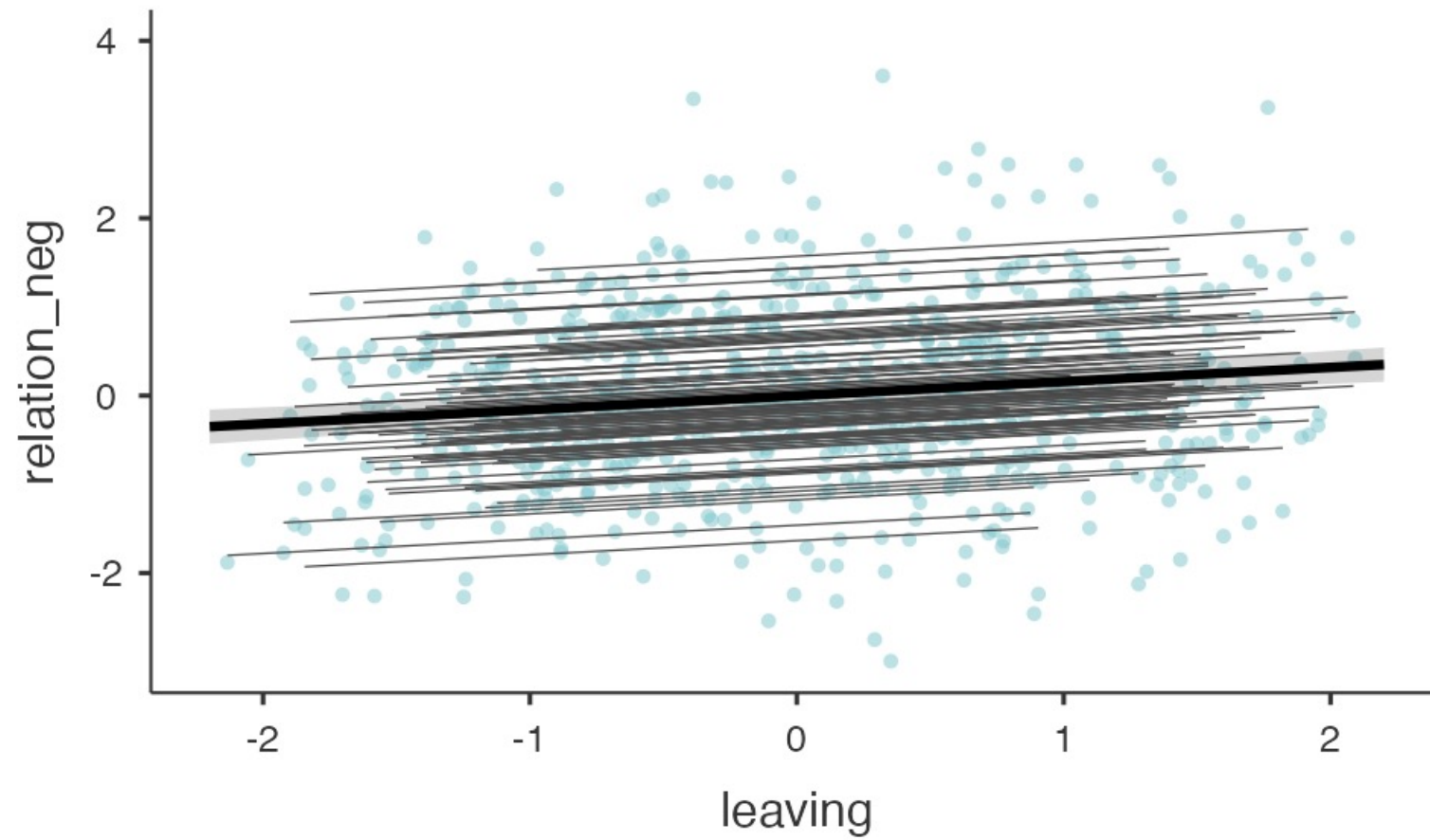
## Effects Plots



## Effects Plots







## Ejercicio 1

- Abrir base de datos "Ejemplo 2.omv".
- Definir el modelo de intercepto aleatorio para predecir los aciertos teniendo en cuenta la variabilidad inter-sujeto.
- Añadimos la condición como predictor. **¿Cómo interpretemos el efecto?**



¿Qué es *realmente* un efecto aleatorio?

## Efectos fijos vs efectos aleatorios

**Regla n° 1:** No preguntes a tu economista de confianza.

**Regla n°2:** No preguntes a tu colega que trabaja en un área experimental.

**Regla n°3:** Tampoco los metodólogos somos de mucha ayuda aquí.



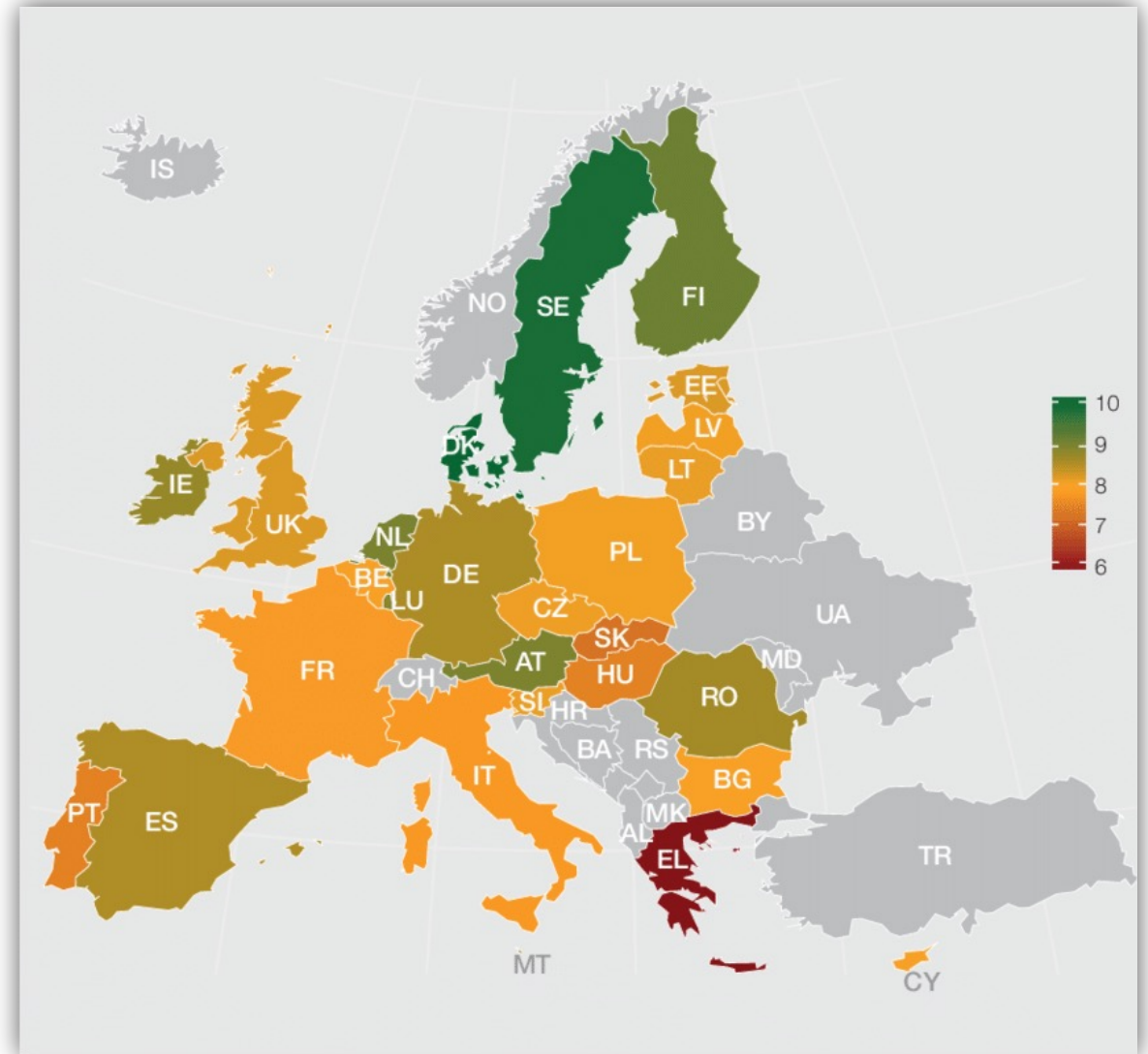
## Visión tradicional

### **Efecto fijo:**

- Niveles conocidos
- Sin distribución de referencia
- No generalizable
- Ej: tratamiento, género, etc.

### **Efecto aleatorio:**

- Niveles desconocidos
- Con distribución de referencia
- Generalizables
- Ej. Individuos, tiempos, etc.



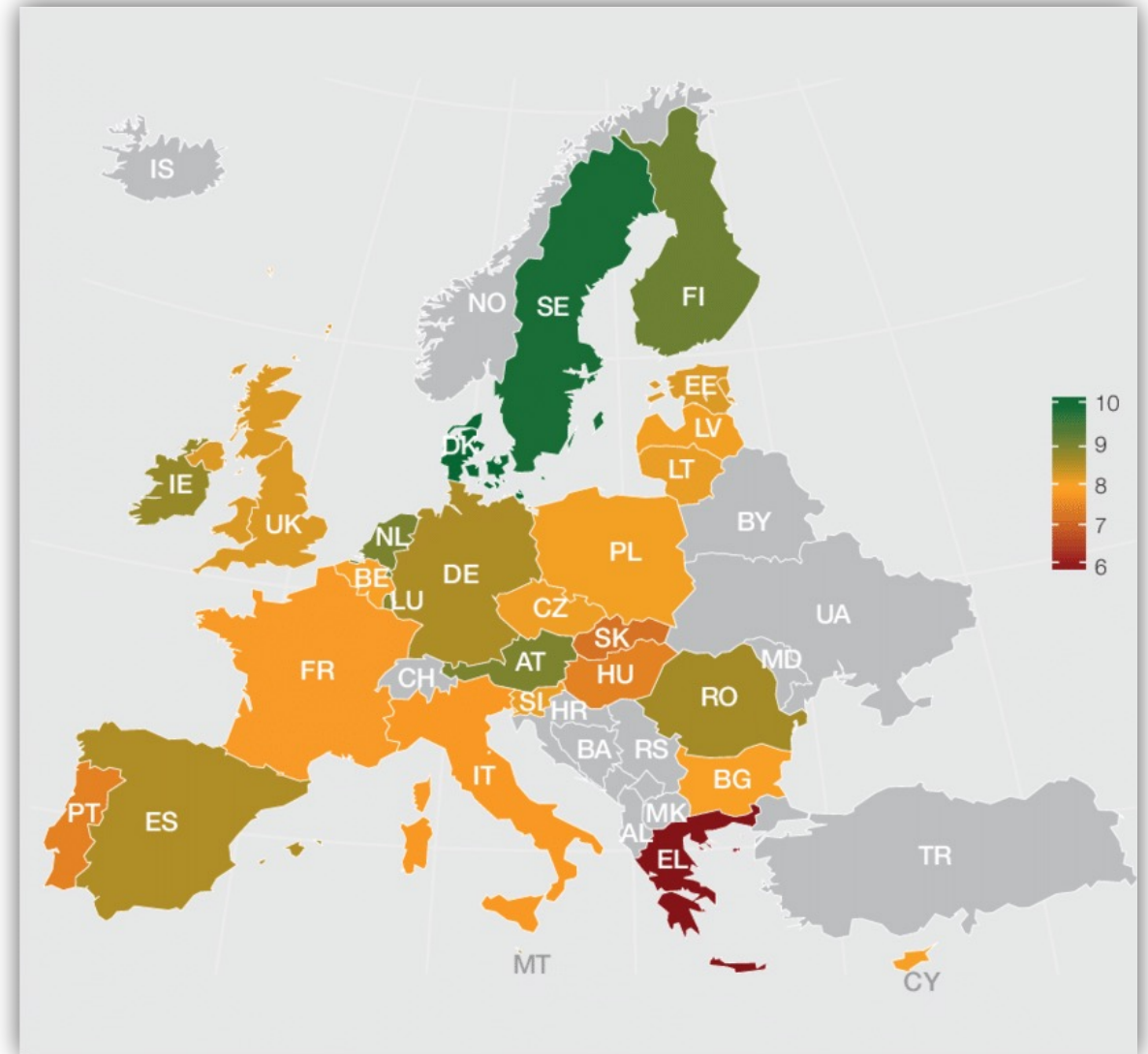
## Visión moderna

### Efecto fijo:

- Sin distribución de referencia.
- Efectos común a todos los elementos.

### Efecto aleatorio:

- Con distribución de referencia.
- Efecto variable a través de las unidades.
- Aportan **regularización**.



Here are the five definitions I've seen:

(1) Fixed effects are constant across individuals, and random effects vary.

Kreft and De Leeuw (1998) thus distinguish between fixed and random coefficients.

(2) Effects are fixed if they are interesting in themselves or random if there is interest in the underlying population. Searle, Casella, and McCulloch (1992, Section 1.4) explore this distinction in depth.

(3) “When a sample exhausts the population, the corresponding variable is *fixed*; when the sample is a small (i.e., negligible) part of the population the corresponding variable is *random*.” (Green and Tukey, 1960)

(4) “If an effect is assumed to be a realized value of a random variable, it is called a random effect.” (LaMotte, 1983)

(5) Fixed effects are estimated using least squares (or, more generally, maximum likelihood) and random effects are estimated with shrinkage (“linear unbiased prediction” in the terminology of Robinson, 1991). This definition is standard in the multilevel modeling literature (see, for example, Snijders and Bosker, 1999, Section 4.2) and in econometrics.

[https://statmodeling.stat.columbia.edu/2005/01/25/why\\_i\\_dont\\_use/](https://statmodeling.stat.columbia.edu/2005/01/25/why_i_dont_use/)



# Random Effects Old and New

James S. Hodges<sup>1\*</sup>, Murray K. Clayton<sup>2</sup>

<sup>1</sup>Division of Biostatistics, University of Minnesota, Minneapolis, Minnesota USA 55414

<sup>2</sup>Department of Statistics, University of Wisconsin-Madison, Madison, Wisconsin USA 53706

*\*email:* hodge003@umn.edu

February 2, 2011

All such new-style random effects can be understood as formal devices to facilitate smoothing or shrinkage, interpreting those terms broadly.

<http://www.biostat.umn.edu/~hodges/PubH8492/Hodges-ClaytonREONsubToStatSci.pdf>





*The term “shrinkage” may have negative connotations. John Tukey preferred to refer to the process as the estimates for individual subjects “borrowing strength” from each other. This is a fundamental difference in the models underlying mixed-effects models versus strictly fixed effects models. In a mixed-effects model we assume that the levels of a grouping factor are a selection from a population and, as a result, can be expected to share characteristics to some degree. Consequently, the predictions from a mixed-effects model are attenuated relative to those from strictly fixed-effects models.*

<https://www.tjmahr.com/plotting-partial-pooling-in-mixed-effects-models/>



## No pooling

- Modelamos cada individuo independientemente.
- Efectos fijos (individuales)
- No se comparte información.
- **Sobreajuste.**

## Partial pooling

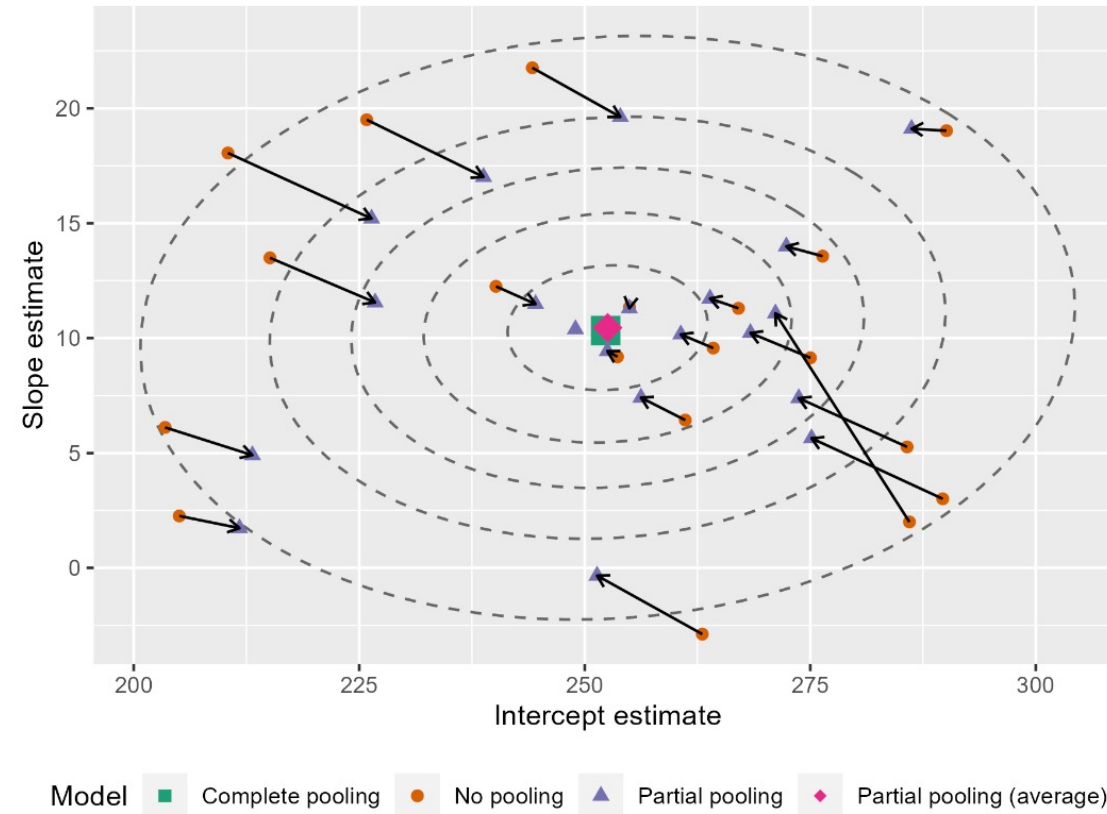
- Modelamos cada individuo como  $\sim N(0, \sigma)$ .
- Efectos aleatorios.
- Se comparte información.
- **Ajuste adecuado\*.**

## Complete pooling

- Estimación común a todos los individuos.
- Efectos fijos (comunes).
- Se comparte *demasiada información.*
- **Infrajuste\*.**

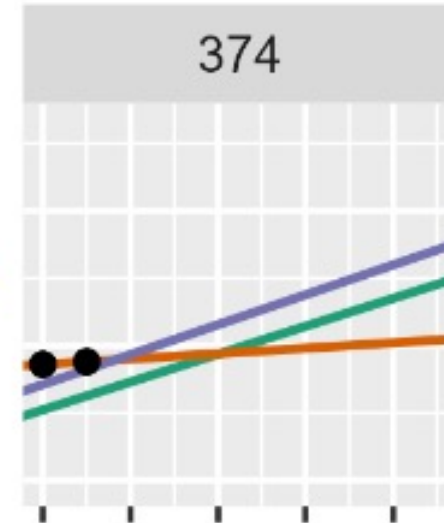
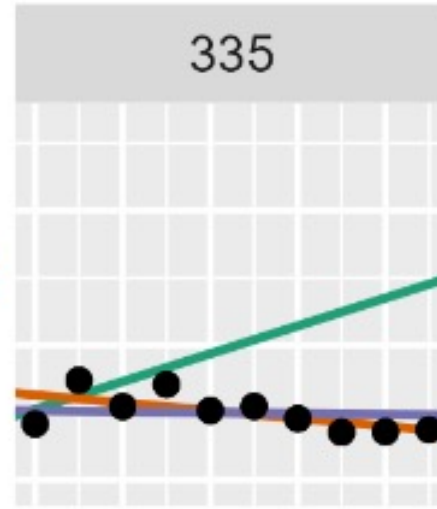
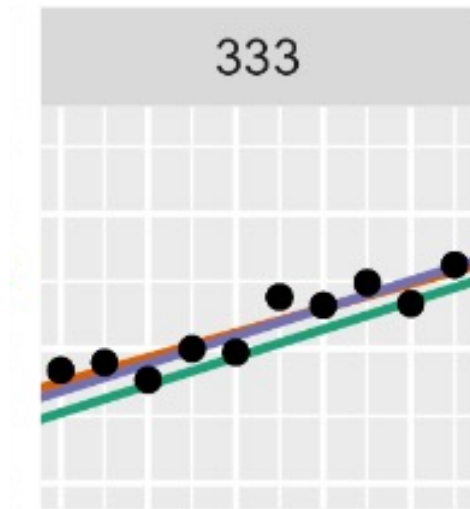


Topographic map of regression parameters



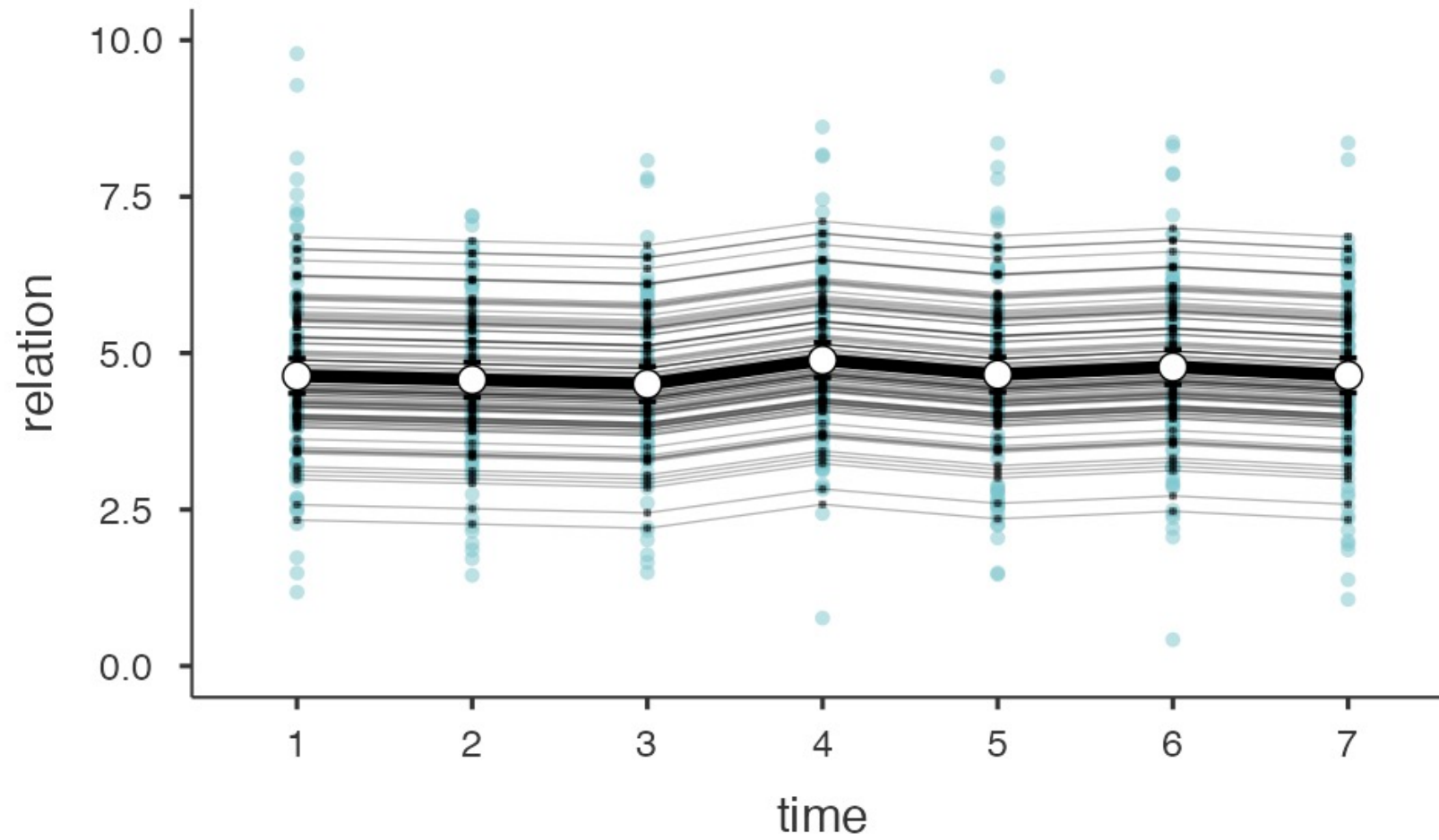
<https://www.tjmahr.com/plotting-partial-pooling-in-mixed-effects-models/>

Model  Complete pooling  No pooling  Partial pooling



<https://shiny.psy.gla.ac.uk/Dale/multilevel/>

## Effects Plots



# Thank you for your attention



eduardo.garciag@uam.es



@edu\_gargar



edugargar

Design by Dr. Ruggeri (Columbia University)

