

# The A/B Effect

# Paper

Michelle N. Meyer, Patrick R. Heck, Geoffrey S. Holtzman, Stephen M. Anderson, William Cai, Duncan J. Watts, and Christopher F. Chabris (2019) Objecting to experiments that compare two unobjectionable policies or treatments

<https://www.pnas.org/content/116/22/10723>

Antes de hablar del paper



### A/B Testing (o randomized controlled trial)

- Dividir una población en dos o más grupos aleatoriamente
- Probar dos estímulos distintos y ver el resultado



A veces la gente miente (o sobreestima su capacidad).

Es necesario exponerles a una situación para ver su reacción.

## A/B testing



Si hay que exponerles a la situación,  
hay que pedir su consentimiento.

A veces obtenerlo afecta el  
resultado.

# Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research, Report of the National Commission for the Protection of



Invitar a participar de una investigación pero que los detalles no se les van a revelar hasta que concluya.

Requisitos:

- que los riesgos para los individuos no sea más que mínimos
- que haya un plan para informar a los sujetos cuando el experimento haya concluido
- no ocultar riesgos y responder cualquier pregunta que hagan las personas.

# Ejemplos



**facebook**

# News Feed



- Primera versión: knobs manuales (más fotos, menos historias de plataforma)
- 2007: modelo basado en afinidad, peso y decaimiento
- 2010: Machine Learning con muchos parámetros

## News Feed



¿Qué efecto genera?

Dos opciones:

- Contagio
- Envidia

The New York Times

## *Facebook Tinkers With Users' Emotions in News Feed Experiment, Stirring Outcry*

By Vindu Goel

June 29, 2014



Facebook revealed that it had altered the news feeds of over half a million users in its study.

Karen Bleier/Agence France-Presse — Getty Images

To Facebook, we are all lab rats.

**okc**

## Match Rating



Trata de estimar cuán compatibles son dos personas.

The New York Times

## *OKCupid Plays With Love in User Experiments*



Mingling at an event in Manhattan sponsored by OKCupid, which on Monday published the results of three experiments. Yana Paskova for The New York Times

By Molly Wood

July 28, 2014



Love is not blind, as it turns out. But opposites attract when people think they are similar.





## Encouraging messages



Software que enseña programación.

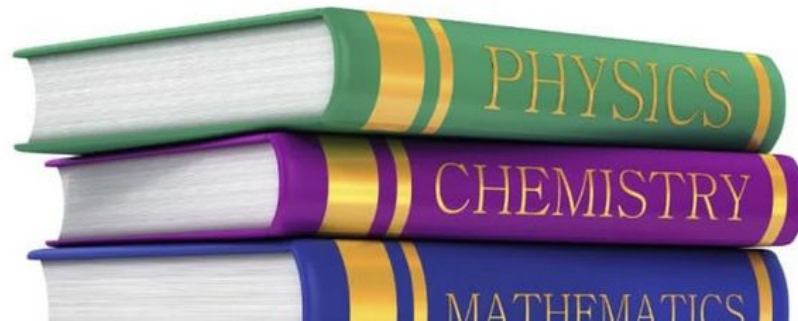
A: no agregar

B: "The most successful programmers spend hours practicing their skills behind the scenes."

C: "Some students tried this question 26 times! Don't worry if it takes you a few tries to get it right."

**The Washington Post**  
*Democracy Dies in Darkness*

# Pearson conducts experiment on thousands of college students without their knowledge



So

Objecting to experiments that  
compare two  
unobjectionable policies or  
treatments

## Estudio 1

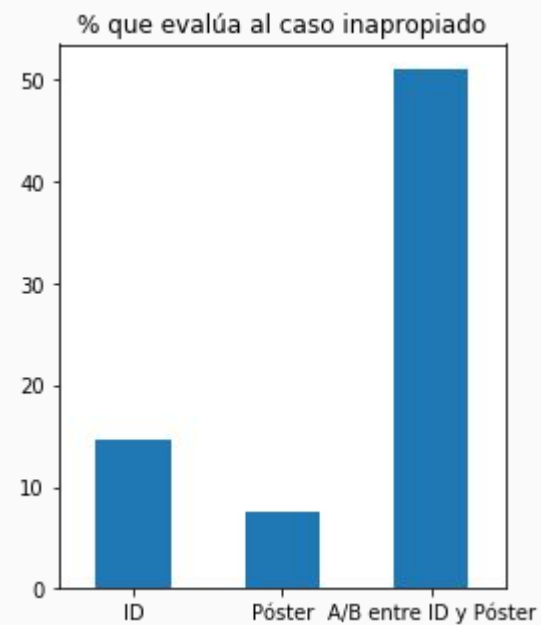


Reducir infecciones al hacer un procedimiento.

Opciones:

- Checklist en ID
- Checklist en poster
- A/B test entre ID y Poster

## Estudio 1 - Resultados



## Generación de hipótesis



¿Por qué puede pasar que haya gente que le parezca inapropiado el experimento?

## Hipótesis 1



Considerar que una de las dos opciones es mejor que la otra.

El experimento no es necesario.

Expone a los pacientes de la opción “inferior” a riesgos innecesarios.



## Estudio 4

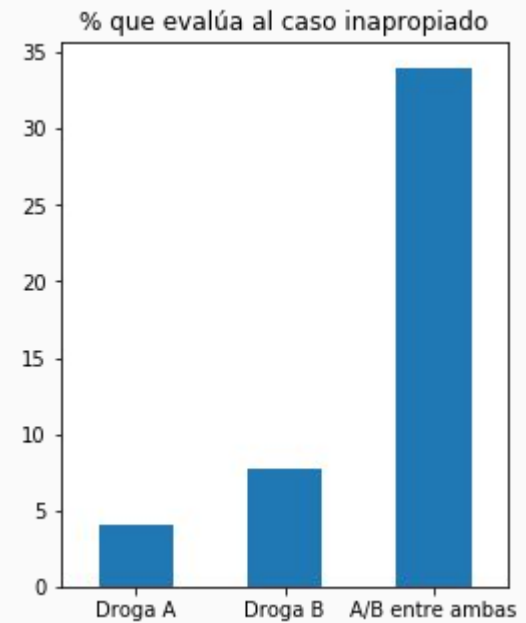


Dadas dos drogas A y B que están aprobadas y sirven no se sabe que una sea mejor que la otra.

Se ofrecen tres opciones: droga A, droga B y un experimento que las compare.



## Estudio 4



## Hipótesis 2



Aversión a la aleatoriedad. Algunas personas recibirían un tratamiento inferior y eso es injusto.

## Estudio 5

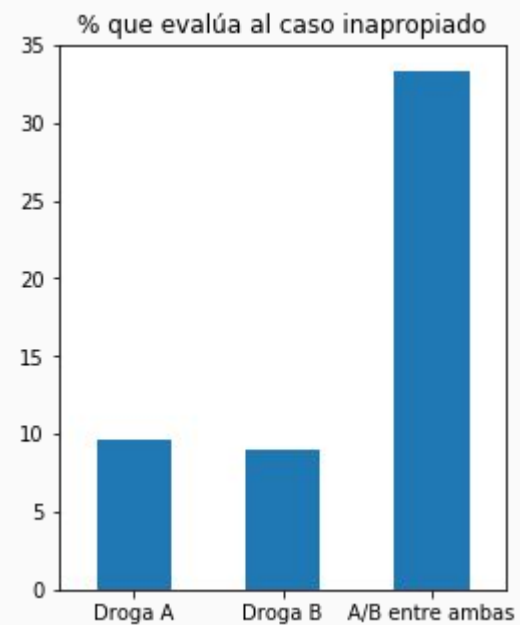


Nueva prueba.

En una clínica algunos doctores prefieren la droga A y algunos la droga B. El Doctor Jones elige entre:

- Caso A: Droga A
- Caso B: Droga B
- Caso A/B: Experimento comparando los resultados.

## Estudio 5



## Hipótesis 3



Falta de consentimiento.

No se necesita consentimiento para agregar una checklist para reducir errores.

Ambas drogas eran aceptables sin consentimiento.

## Hipótesis 4



### Motivación vil

Aunque se haya explicado el porqué del experimento quizás hay gente que considera que se hacen por fines no altruistas.



## Hipótesis 5



Creencia de que los expertos ya deberían saber qué es mejor.

Esto también hace menos aceptables las pruebas y más aceptables las soluciones no probadas.

## Hipótesis 6



Aversión a la experimentación

("guinea pigs", "playing with lives",  
"experimenting")



## Hipótesis 7



Nivel de educación

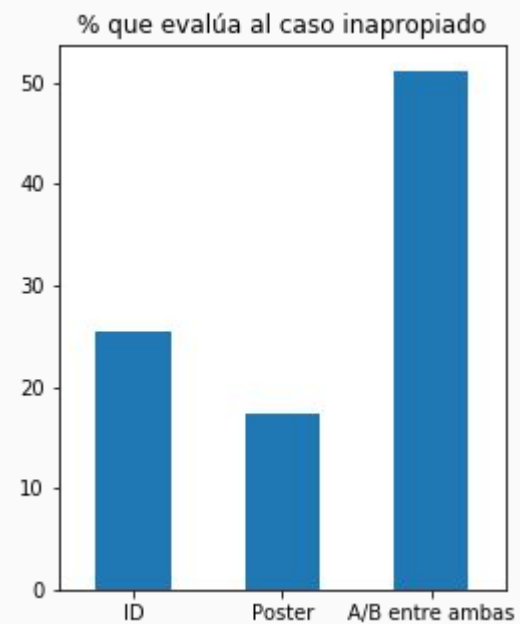
No hay una correlación entre el nivel educativo o carrera de la gente que respondió y los resultados.

## Estudio 6

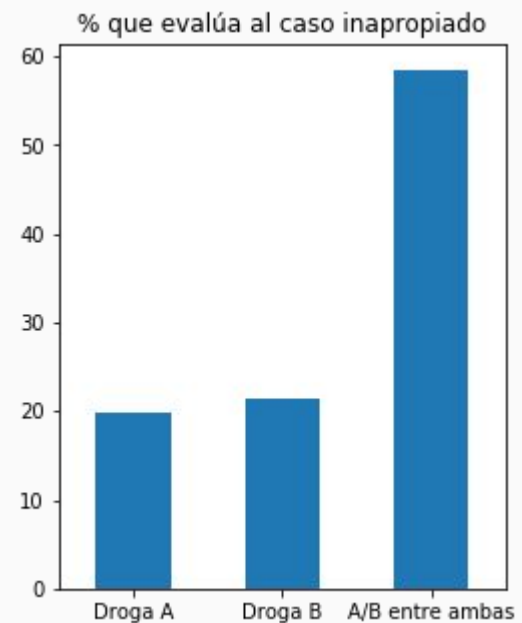


Replicar estudio 1 (checklist en ID vs poster) y 5 (drogas A vs B) con profesionales de la salud.

## Estudio 6



## Estudio 6



# Conclusiones

# Resumen

*When neither of two policies is objectionable or perceived as clearly superior, an A/B test comparing them should not be seen as more morally problematic than a unilateral decision to implement either untested policy.*

# Conclusiones

Las personas que estén en contra de los experimentos pueden negarse a participar, sesgando la muestra de los experimentos y perjudicando.

# Conclusiones

Este rechazo a la experimentación puede hacer que los políticos tomen decisiones no probadas en vez de evaluarlas primero.

Ya sea por ellos caer en esta ilusión o por pensar que sus votantes pueden pensarlo.



¿Preguntas?



¡Gracias!



# Links

<https://github.com/seppo0010/the-a-b-effect>

<https://www.pnas.org/content/116/22/10723>