

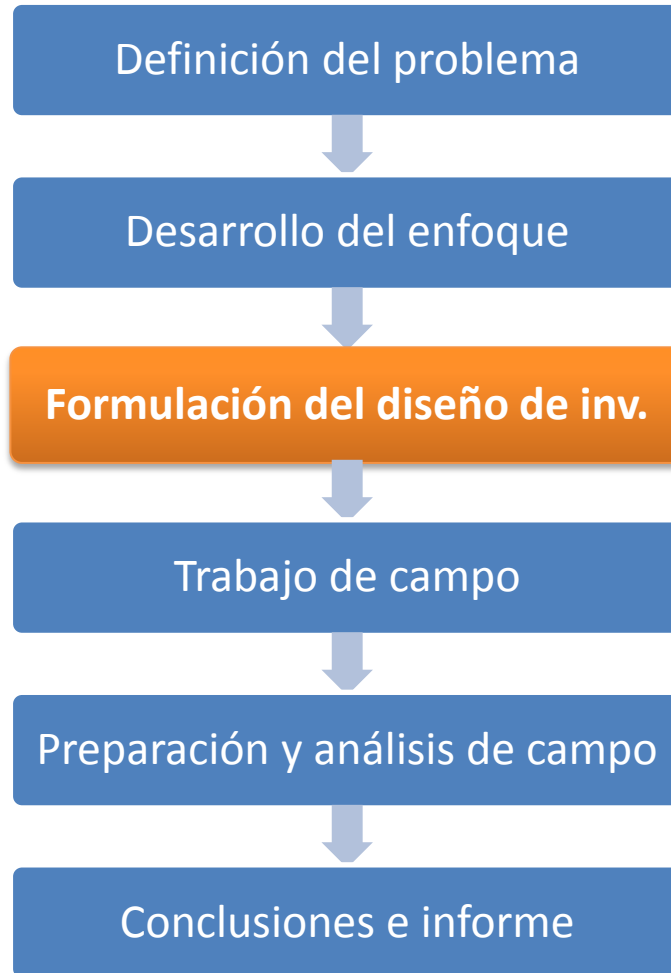
IN5625 - INVESTIGACIÓN DE MERCADOS

Diseño de la investigación – Investigación causal

André Carboni

Semestre primavera 2013

acarboni@ing.uchile.cl



A close-up photograph of two glass test tubes being held by hands. The tube on the left contains a red liquid and has a scale from 0 to 14 ml. The tube on the right contains a green liquid and has a scale from 0 to 25 ml. The tubes are tilted and their mouths are touching, suggesting they are being mixed. The background is a blurred laboratory with shelves containing various bottles and containers.

INVESTIGACIÓN CAUSAL



- **Investigación causal:** tipo de investigación concluyente donde el objetivo principal es obtener evidencias respecto a las relaciones de causa y efecto.
- **Causalidad:** ocurre cuando la presencia de la variable X afecta la probabilidad de presencia de la variable Y .

Significado común	Significado científico
X es la única causa de Y	X es una de las causas de Y
X siempre debe llevar a Y	X afecta la probabilidad de Y
Es posible mostrar que X es causa de Y	A lo mas podemos inferir que X es causa de Y

Se deben satisfacer 3 condiciones (necesarias pero no suficientes) antes de hacer inferencia causal o de suponer causalidad:

1. Variación concomitante:

- Los efectos de las variables X e Y varían juntos.
- Cuantitativamente este efecto se denomina **correlación**.
- Ej: Encuesta a 1.000 personas, se obtiene que 73% de los encuestados con “alto nivel académico” tienen un alto nivel de compras de ropa de moda, mientras que sólo el 64% de los de “bajo nivel académico” tienen un alto nivel de compras.
 - ¿Se concluye que alto nivel académico causa compra elevada de ropa de moda? R: **¡No!** La evidencia hace más defendible la hipótesis, pero no la demuestra (podría, por ejemplo, deberse a otro factor como el ingreso).

2. Orden temporal de la ocurrencia de las variables:

- El evento causal debe ocurrir **antes o, a lo más, simultáneamente** que el evento efecto. No puede ocurrir después.

3. Ausencia de otros factores causales posibles:

- En toda investigación causal se debe buscar minimizar el efecto de otras variables, aislando la causa y efecto del medio ambiente.



- **Diseño experimental:** conjunto de procedimientos que especifican:
 - **Unidades de prueba:** entidades respecto de las cuales se analiza la relación causal.
 - **Variables independientes:** conjuntos de elementos manipulables cuyos efectos se miden en una variable dependiente.

- **Variable dependiente:** elemento sobre el cual se mide el efecto de las variables independientes en las unidades de prueba.
- **Variables extrínsecas:** elementos que afectan las unidades de prueba, pero que no han sido considerados como variables independientes. **Se deben controlar.**

Ej:



- **Experimento:** cuando el investigador manipula una o mas variables independientes y mide el efecto que tiene sobre la variable dependiente.

- **X** = Exposición de un grupo a una variable independiente, tratamiento o evento cuyos efectos deben determinarse.
- **O** = Proceso de observación o medida de la variable dependiente en las unidades de prueba.
- **R** = Asignación aleatoria de unidades de prueba a grupos o tratamientos separados.
- El tiempo se mueve de izquierda a derecha.
- Cada fila corresponde a un grupo de trato específico.
- Cada columna corresponde a eventos simultáneos

- Ejemplos:

X	O ₁	O ₂
---	----------------	----------------

Significa que un grupo de unidades de prueba se expuso a la variable independiente (o “de tratamiento”) (X) y que la respuesta se midió en dos momentos diferentes O₁ y O₂

R	X ₁	O ₁
R	X ₂	O ₂

Significa que dos grupos de unidades de prueba fueron asignados al azar a dos grupos de tratamiento diferentes y que la variable dependiente se midió simultáneamente en ambos grupos.

- **Validez interna:** medida de la precisión de un experimento. Mide si la manipulación de las variables independientes o tratamientos, provoca en realidad los efectos en la variable dependiente.
- **Validez externa:** determinación de si las relaciones de causa y efecto descubiertas en el experimento pueden generalizarse.

Existe un trade-off entre una y otra validez en el diseño del experimento: Se puede realizar un experimento en un ambiente artificial para controlar variables extrañas. Esto mejora validez interna, pero reduce la externa.

- **Historia (H):** El medio ambiente que influencia a las unidades de prueba. Eventos que suceden al mismo tiempo y que son externos al experimento. Considere el siguiente experimento:

O_1 X_1 O_2

Donde O_1 y O_2 son ventas de una tienda y X_1 es una campaña promocional. Supongamos que el experimento reveló que no había diferencia entre O_2 y O_1 . ¿Podemos decir que la campaña fue infructuosa? ➔ ¡NO! Economía, desempleo, ...

- **Madurez (MA):** Cambios en las unidades de prueba causadas por el tiempo (y **no** por las variables independientes). Ej: Cansancio, envejecimiento, aburrimiento, pérdida de interés.
- **Efecto de la prueba:** Cambios que pueda producir el mismo experimento. Por lo general son los efectos sobre el experimento de hacer una medición de la variable dependiente antes y después de la presentación del tratamiento.
 - **Efecto principal de la prueba (EP):** Ocurre cuando una observación anterior afecta a una posterior.
 - **Efecto interactivo de la prueba (EI):** Ocurre cuando una medición previa afecta la respuesta de las unidades de prueba a la variable independiente.
 - Ej:



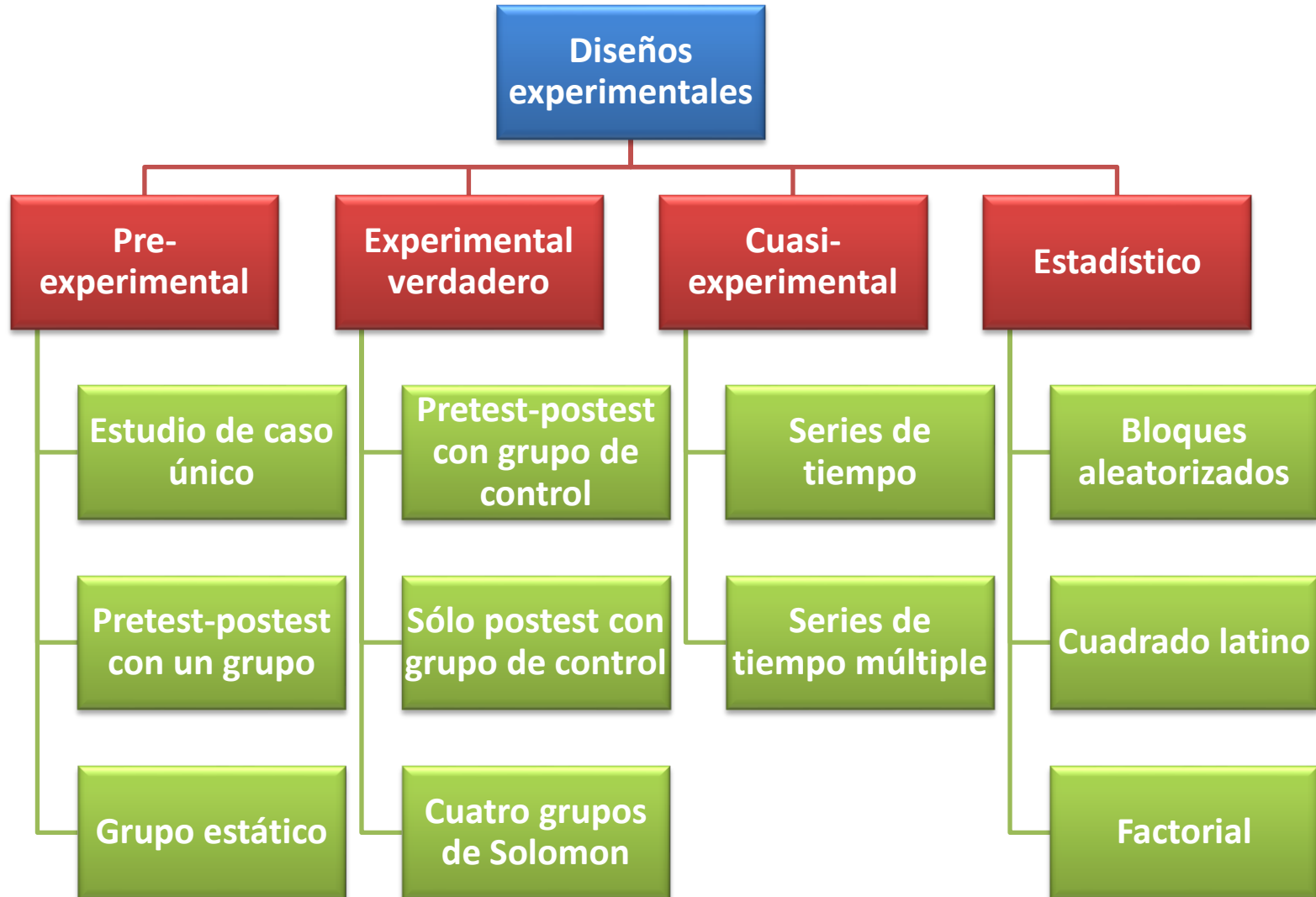
- **Instrumentación (I):** Cambios en los instrumentos de medición, en los observadores o en las puntuaciones mismas. Ej: Experimento donde se miden las ventas en \$ antes y después de un tratamiento (X), y hay un cambio de precios no experimental.
- **Sesgo de selección (SS):** Elección de muestras no uniformes/inadecuadas. Ocurre cuando la selección de unidades de prueba resulta en grupos de tratamiento que difieren en la variable dependiente, antes de ser expuestos a la condición de tratamiento (X).
- **Mortalidad (MO):** Pérdida de unidades de prueba mientras transcurre el experimento. Ej: Unidades de prueba que se niegan a continuar con el experimento.

- Las variables externas representan explicaciones alternativas de los resultados experimentales. Representan grave amenaza a la validez interna y externa → Hay que controlarlas.
1. **Aleatorización:** Asignación al azar de las unidades de prueba a los grupos experimentales mediante uso de números aleatorios. Las condiciones de tratamiento también se asignan al azar → ¡Los factores externos se ven en igual proporción en cada grupo!
 2. **Pareamiento:** Supone la comparación de unidades de prueba en un conjunto de variables antecedentes clave, antes de asignarlos a las condiciones de tratamiento.

- **Control estadístico:** implica medir las variables externas y ajustar sus efectos mediante análisis estadístico. Ej: ANCOVA (Análisis covarianza). En este se eliminan los efectos de la variable externa sobre la variable dependiente.
- **Diseño de control:** Supone el uso de experimentos diseñados para controlar variables extrañas específicas.

- **Diseños pre-experimentales:** no utilizan procedimientos de aleatorización para controlar los factores externos.
- **Diseños experimentales verdaderos:** Se asigna al azar las unidades de prueba y los tratamientos a los grupos experimentales.
- **Diseños cuasi-experimentales:** Resultan cuando el investigador es incapaz de manipular por completo la programación o asignación de los tratamientos a las unidades de prueba, pero aplica algo del aparato de la experimentación verdadera.
- **Diseño estadístico:** Serie de experimentos básicos que permiten el control y análisis estadístico de las variables externas.

Clasificación de los diseños experimentales



1. Estudio de caso único:

X

O₁

- Un solo grupo de unidades de prueba, sin grupo de control.
- Problema: No hay base para comparar O₁.
- Poca validez interna, ya que pueden afectar muchas variables externas.
- Solo para investigación exploratoria.

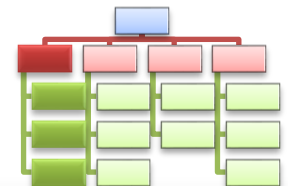
2. Diseño de pretest y posttest con un grupo:

O₁

X

O₂

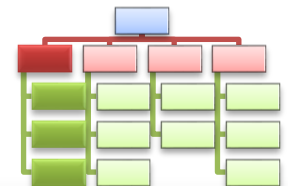
- Un solo grupo de unidades de prueba, dos mediciones, sin grupo de control
- El efecto de X se mide como $(O_2 - O_1)$, aunque su validez es cuestionable si no se controlan variables externas.
- Solo cuando variables extrínsecas muy controladas.



3. Diseño de grupo estático:

(GE)	X	O ₁
(GC)		O ₂

- Dos grupos: “experimental” y “control”.
- Grupo experimental se expone al tratamiento (X), el de control **no**.
- Efecto = $(O_1 - O_2)$.
- Problemas: Selección y mortalidad.
- Solo cuando se presume que la selección de la muestra no es tan relevantes

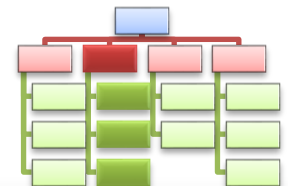


- **Diseño de pretest-posttest con grupo de control:**

(GE) R O_1 X O_2

(GC) R O_3 O_4

- Unidades de prueba asignadas al azar entre GE y GC.
- Efecto en el grupo experimental: $(O_2 - O_1) - (O_4 - O_3)$
- Este diseño controla la mayoría de las variables externas.



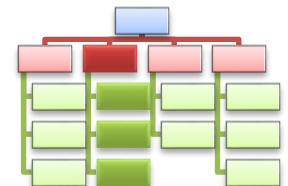
- **Diseño de solo posttest con grupo estático**

(GE)	R	X	O_1
(GC)	R		O_2

- El efecto del tratamiento: $(O_1 - O_2)$
- Muy sencillo de poner en práctica. Susceptible de sesgo de selección y mortalidad.

- **Diseño de cuatro grupos de Solomon**

- Mizerki, Allison & Calvert, “A controlled field study of corrective advertising using multiple exposure and a commercial medium”, Journal of marketing research (1980).



- **Diseño de series de tiempo:**

$O_1 \quad O_2 \quad O_3 \quad X \quad O_4 \quad O_5 \quad O_6$

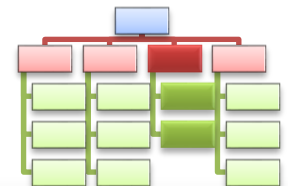
- Debilidad: Incapacidad para controlar la historia.
- Ej: Panel de prueba, se muestra un comercial cierta cantidad de veces en un mes. No podemos saber si los individuos lo vieron o cuando. Se examinan las compras.
- Útil cuando los experimentos pueden reportar efectos retardados o con distintos niveles de influencia a través del tiempo.

- **Diseño de series de tiempo múltiples:**

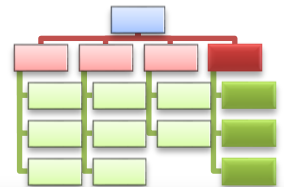
(GE) $O_1 \quad O_2 \quad O_3 \quad X \quad O_4 \quad O_5 \quad O_6$

(GC) $O_7 \quad O_8 \quad O_9 \quad O_{10} \quad O_{11} \quad O_{12}$

- Se disminuye la posibilidad de error por la influencia de variables extrínsecas.



- **Diseño de bloques aleatorios:**
 - las unidades de prueba se dividen en bloques con base en una variable externa que se presume influya en la variable dependiente.
 - Ej: Hacer un mismo experimento para cada tramo de ingresos para estudiar un efecto en el cambio del precio de un producto.



Ejemplo bloques aleatorios

	Comercial A	Comercial B	Comercial C
Alto conocimiento del producto	A	B	C
Intermedio conocimiento del producto	A	B	C
Bajo y ninguno conocimiento del producto	A	B	C

A, B y C son los 3 comerciales de prueba:

- A: ningún humor.
- B: cierto humor.
- C: mucho humor.

- **Diseño de cuadro latino:**
 - Permite el control (estadístico) de dos variables externas que no interactúan, además de la manipulación de la variable independiente.
 - La variable independiente y cada variable externa se divide en igual número de niveles
 - La regla de asignación es que cada nivel de variable independiente se asignan a la tabla de modo que aparezca una y solo una vez en cada fila y columna.

Ejemplo cuadro latino

	Alto interés en el producto	Intermedio interés en el producto	Bajo interés en el producto
Alto conocimiento del producto	B	A	C
Intermedio conocimiento del producto	C	B	A
Bajo y ninguno conocimiento del producto	A	C	B

A, B y C son los 3 comerciales de prueba:

- A: ningún humor.
- B: cierto humor.
- C: mucho humor.

- **Diseño factorial:** se utiliza para medir los efectos de dos o mas variables independientes en diversos niveles y permite la interacción entre las variables.
- Ejemplo:

	Ningún humor	Humor medio	Humor alto
Baja información entregada	A	B	C
Media información entregada	D	E	F
Alta información entregada	G	H	I

Experimentos de laboratorio v/s de campo

Factor	Laboratorio	Campo
Ambiente	Artificial	Realista
Control	Alto	Bajo
Error reactivo	Alto	Bajo
Posibilidad de plantear escenario	Alto	Bajo
Validez interna	Alta	Baja
Validez externa	Baja	Alta
Tiempo	Corto	Largo
Nº de unidades	Pequeño	Grande
Fácil implementar	Alta	Baja
Costo	Bajo	Alto