

# MÉTODOS Y DISEÑOS DE INVESTIGACIÓN

Profesor: Alfonso Pitarque

[www.uv.es/pitarque](http://www.uv.es/pitarque)

## TEMARIO

Tema 1: EL DISEÑO DE UNA INVESTIGACIÓN EN PSICOLOGIA

Tema 2: DISEÑOS EXPERIMENTALES.

Tema 3: ANÁLISIS DE LOS DISEÑOS EXPERIMENTALES.

Tema 4: INTRODUCCION A LA INVESTIGACIÓN CUASI EXPERIMENTAL.

Tema 5: INTRODUCCION A LA INVESTIGACIÓN POR ENCUESTAS.

Tema 6: INTRODUCCION A LA INVESTIGACIÓN OBSERVACIONAL.

## BIBLIOGRAFÍA.

→ Gambara, H. (2002). *Métodos de Investigación en psicología y Educación. Cuaderno de prácticas (3ª edición)*. McGraw-Hill: Madrid.

→ León, O.G. y Montero, I. (2003). *Métodos de Investigación en psicología y Educación (3ª edición)*. McGraw-Hill: Madrid.

→ Meltzoff, J. (2000). *Crítica a la investigación. Psicología y campos afines*. Alianza Editorial. Madrid. (pps-195-321).

# TEMA 1. EL DISEÑO DE UNA INVESTIGACIÓN EN PSICOLOGIA.

## 1.1. INTRODUCCION.

El hombre siempre ha necesitado conocer y controlar su entorno para vivir más y mejor. La necesidad incentiva la curiosidad, base del conocimiento.

En los primeros tiempos ese conocimiento lo daban las distintas supercherías y religiones. Pero con los árabes (Avicena, ...) y sobre todo a partir del Renacimiento el hombre comienza a creer que es capaz de buscar soluciones por sí solo (p.e. Da Vinci, Servet, Vesalio, Copérnico,...).

Hoy la ciencia es el principal modo de adquirir conocimiento (pero no el único: p.e. refranes, artes, religión, magia,...) para control del entorno y garantía de supervivencia de la especie humana (Medicina, Ingenierías, Informática, etc.).

Definir la Ciencia es difícil. Lo que define a una ciencia no es su objeto de estudio si no su método de trabajo: el método científico o hipotético-deductivo.

El método científico se caracteriza porque sigue una serie de pasos secuenciales:

- 1) Plantea un problema de investigación
- 2) Deriva unas hipótesis contrastables en términos operativos (si ... entonces ...)
- 3) Contrasta dichas hipótesis empíricamente (es decir a través de la medición cuantitativa = objetiva) utilizando diseños de investigación apropiados
- 4) Saca conclusiones que avalarán o refutarán teorías

De este modo el conocimiento que nos aporta dicho método es:

- 1) Objetivo (no subjetivo)
- 2) Determinista (todo tiene una causa y ésta siempre es cognoscible)
- 3) General (busca leyes generales del conocimiento)
- 4) Contrastable empíricamente y por tanto refutable y relativo (abierto siempre a la revisión, es decir replicable)
- 5) Parsimoniso

## 1.2. METODOS DE INVESTIGACION EN PSICOLOGIA.

La Psicología ganó el estatus de Ciencia sólo cuando demostró que era posible cuantificar y medir objetivamente lo mental y adoptó el método científico en su forma de trabajo:

- Fechner (1860), Wundt (1879), Ebbinghaus (1885), Wertheimer (1912) en Alemania,
- Stanley Hall (1883), William James (1890), Thorndike (1898) y Watson (1913) en USA,
- Binet y Simon (1905) en Francia
- Pavlov (1906) en Rusia.

Desde entonces se ha hablado siempre de dos tradiciones o enfoques en la investigación psicológica:

- a) el enfoque correlacional (o psicométrico) iniciado por Binet y Simon y después Wechsler, cuyo objetivo es construir tests y cuestionarios para medir lo psíquico haciendo especial hincapié en las diferencias individuales y en el análisis de las relaciones entre las variables medidas (correlación, regresión, análisis factorial, etc.)
- b) el enfoque experimental basado en la búsqueda de las leyes generales de la conducta mediante la experimentación de laboratorio y el establecimiento de relaciones causales entre variables, priorizando técnicas estadísticas inferenciales como las pruebas t y el análisis de la varianza.

## Inciso termonológico:

En el diseño de una investigación debemos distinguir entre tres tipos principales de variables:

- Variables independientes (VI; también llamadas predictoras) y sus niveles (o condiciones o tratamientos). Son las que se manipulan, bien directamente por el investigador (manipulación directa) o bien simplemente seleccionando sus niveles de la naturaleza por el investigador (manipulación por selección). Esta distinción es muy importante porque determina la distinción entre métodos experimentales (= manipulación directa) y cuasi-experimentales (= manipulación por selección).
- Variables dependientes (VD o variables criterio). Son las que medimos para ver si la manipulación anterior es efectiva o no. Podrán medirse en una escala:
  - a. Nominal (=VD cualitativa o categórica)
  - b. Ordinal (= VD semi- cuantitativa)
  - c. Intervalo (= VD cuantitativa)
  - d. Razón (= VD cuantitativa)

- Variables extrañas o contaminantes, que son las que hay que controlar para que no afecten a la VD. Cuando todas las posible variables extrañas están bien controladas se dice que la investigación tiene *validez interna*.

Los distintos métodos de investigación en Psicología se podrían sistematizar en 4 tipos principales en función del mayor o menor grado de control que el investigador tiene sobre las variables implicadas en la investigación:

a) Metodología experimental: Su objetivo es conocer las causas de los fenómenos, los cómo y los por qué de los mismos, es decir, poder establecer relaciones de causa-efecto entre variables. La experimentación es pues la única metodología que permite establecer relaciones de causa-efecto entre variables. Para ello el investigador puede manipular directamente la VI (manipulación directa) y controlar adecuadamente las variables extrañas, generalmente mediante la asignación aleatoria de los sujetos a los tratamientos. Es decir, aquí la validez interna queda garantizada. Utiliza prioritariamente las técnicas estadísticas inferenciales paramétricas (ANOVA) y no paramétricas. La experimentación se puede realizar tanto en laboratorio como en ámbitos naturales. (ejs. Davidoff, pps. 52, 48, 454-5)

b) Metodología cuasi-experimental. Su objetivo es el mismo que el de la experimentación, pero aquí el investigador sólo puede seleccionar muestras en base a características intrínsecas de los sujetos que no son directamente manipulables (p.e. su sexo, edad, nivel educativo,...). Se habla entonces de manipulación por selección. Ello implica que los sujetos no son asignados aleatoriamente a los tratamientos y por tanto nunca podremos tener certeza absoluta de que los cambios que aparecen en la VD son exclusivamente debidos a la manipulación de la VI. La cuasi- experimentación suele utilizarse en ámbitos naturales (la escuela, el lugar de trabajo, etc.). Utiliza las mismas técnicas estadísticas que la experimentación. P.e. calidad de vida en personas que viven solas frente a personas que viven en pareja o con amigos.

c) Metodología de encuestas (o metodología correlacional, psicométrica, selectiva o de tests y cuestionarios). Su objetivo es describir hechos psicológicos y analizar relaciones entre ellos mediante la aplicación de encuestas, tests o cuestionarios a muestras grandes y representativas. Suele utilizar técnicas estadísticas descriptivas (tendencia central, variabilidad, correlaciones, regresión, etc.) (ej. Davidoff, pp. 44)

d) Metodología observacional. Su objetivo describir fenómenos que ocurren en ambientes naturales o de laboratorio sin intervención del investigador. La recogida de datos la suele hacer bien el propio investigador o bien se utilizan vídeos, etc. Se trabaja con muestras pequeñas de sujetos o incluso con un solo sujeto (estudio de casos: p.e. Piaget, Freud, etc.). La investigación observacional suele utilizarse como herramienta para la elaboración de hipótesis que luego pondremos a prueba con alguna de las otras tres metodologías referidas arriba. Las pruebas estadísticas más utilizadas aquí son las tablas de contingencia (generalmente analizadas mediante ji-cuadrado, etc.) (ej. Davidoff, pp. 41)

Algunos autores hablan también de “metodologías cualitativas”: Su denominador común es el rechazo del análisis cuantitativo de datos y de la estadística, priorizando estrategias de investigación no estructuradas y subjetivas. Destacan el estudio de casos (entrevista) y los grupos de discusión. (ej. Davidoff, pp. 46)

Casi todas las Ciencias en sus inicios utilizaron métodos de investigación puramente observacionales y/o correlacionales, para llegar al final a una preponderancia de la metodología experimental/cuasiexperimental. La Psicología no ha sido una excepción.

### 1. 3. EL DISEÑO DE UNA INVESTIGACIÓN.

La aplicación del método hipotético-deductivo en Psicología (y en cualquier otra ciencia), sigue una serie de pasos secuenciales que podrían ser:

- Plantear un problema o pregunta de investigación y documentarse sobre el mismo
- Derivar hipótesis operativas
- Elegir el diseño: VI (niveles y condiciones), VD, estrategias de control de VE, cuántas muestras, decidir el análisis estadístico que haremos
- Decidir sobre el procedimiento y materiales
- Aplicar los tratamientos
- Medir o recoger los datos
- Análisis estadístico de los datos
- Interpretación de los resultados
- Elaboración de un informe escrito

Veamos con mayor detenimiento cada uno de estos puntos:

1) Definición de un problema o una pregunta de investigación. El problema debe de ser:

- Relevante, original y que no se conozca la respuesta (documentarse bien antes)
- estar formulado de forma operativa, es decir, que sea resoluble empíricamente (ejs. cuadros pps. de 29-31 León y Montero)
- que implique usar medios éticos



Como señala Meltzoff (2000) el tipo de pregunta determina en gran medida el diseño de investigación que deberemos utilizar. Así en Psicología las preguntas más comunes suelen ser principalmente de uno de los siguientes 6 tipos:

- Preguntas de descripción: ¿Cómo es x?. ¿Cuáles son sus características? (P.e. ¿cuáles son las características de personalidad de las adolescentes anoréxicas?). Este tipo de preguntas implica metodologías observacionales o de encuestas sobre muestras grandes y representativas. Suele ser necesario contar con un grupo control con el que comparar los resultados.
- Preguntas de composición: ¿Cuáles son los componentes del rasgo x? (P.e. ¿qué factores configuran la autoestima?). Requieren metodología de encuestas sobre muestras grandes y representativas. Muestreo de todos los aspectos del constructo a medir. Suelen analizarse mediante análisis factorial.
- Preguntas de relación (o predicción) entre dos o más variables (p.e. el nivel cultural de los padres y rendimiento académico de los niños están relacionados?). Metodología de encuestas (distintos coeficientes de correlación; técnicas de regresión, etc). Atención: no se garantiza la causalidad.
- Preguntas comparativas entre dos o más grupos preexistentes (p.e. ¿son los hombres más agresivos que las mujeres?). Metodología y diseños cuasi-experimentales: manipulación por selección de variables independientes. Hay que garantizar que los grupos a comparar sean equivalentes en las VE más relevantes. Técnicas de análisis de datos inferenciales paramétricas o no paramétricas: pruebas t, análisis de varianza, etc.

- Preguntas de causalidad (simple): ¿X es la causa de Y (o Y el efecto de X)? ¿Cómo y por qué X determina a Y?. (p.e ¿la psicoterapia X es eficaz para reducir significativamente los pensamientos obsesivo-compulsivos?). Metodología y diseños experimentales unifactoriales: manipulación directa de 1 sola VI, control de variables extrañas generalmente por asignación aleatoria de los sujetos a los tratamientos. Utilización de grupos control (placebo). Técnicas de análisis de datos inferenciales paramétricas o no paramétricas: pruebas t, análisis de varianza (ANOVA) unifactoriales, etc.
  
- Preguntas de causalidad múltiple (interacción): ¿X e Y afectan a Z en ciertas condiciones pero no en otras?. (p.e ¿El fármaco X interactúa con el consumo de alcohol?). Metodología y diseños experimentales o cuasi-experimentales factoriales: manipulación directa de dos o más VI. Utilización de grupos control (placebo). Técnicas de análisis de datos inferenciales paramétricas o no paramétricas: pruebas t, análisis de varianza factoriales, etc.

2) Fase de documentación. Estudio de las teorías y resultados de otros laboratorios que intentan explicar dicho problema.

- En la actualidad la principal fuente de documentación en nuestra ciencia es la consulta de la base de datos informatizada PsycINFO de la APA. Está disponible en la dirección web de nuestra universidad <http://biblioteca.uv.es> seleccionando Recursos electrónicos + Bases de datos + Ciencias Sociales + Psicología + PsycINFO.

- De igual modo la consulta del Social Sciences Citation Index del Web of Knowledge de ISI-Thomson nos permite conocer el índice de impacto de una revista, autor, etc. así como buscar por palabras clave (como en PsycINFO). Está disponible en la dirección web de nuestra universidad <http://biblioteca.uv.es> seleccionando Recursos electrónicos + Bases de datos + Multidisciplinarios + Web of Knowledge (WOK) + Acceso a la WOK (cuadro a la izquierda) y:

- Web of Science (activar sólo SSCI) + *General search* para buscar sobre un tópico, autor o revista (pulsar submit marks + marked lists para grabar o enviar por e-mail)

- Web of Science (activar sólo SSCI) + *Cited reference search* para saber qué trabajos citan a un autor, artículo, etc.

- Para buscar el índice de impacto de una revista: ISI Journal Citation Report (+go) + JCR Social Sciences Edition 2007 + View a group of journal by subject category + psychology (applied, biological, clinical, etc.)

- Consulta de revistas de recopilación: Annual Review of Psychology (<http://www.annualreviews.org/>), Psychological Review, Psychological Bulletin, (<http://www.apa.org/journals>).

- En castellano: consulta de LATINDEX ([www.latindex.org](http://www.latindex.org)) con enlaces a revistas electrónicas, y de la base IN-Recs para ver el índice de impacto de revistas españolas (<http://ec3.ugr.es/in-recs/>)

- Consulta de Google Scholar: <http://scholar.google.com>

- Consulta de Google Print: <http://books.google.com>

3) Deducción de hipótesis contrastables (operativas) lo más concisas posible. Mejor en formato SI ...  
ENTONCES ...

4) Elección del diseño: Es un parte muy delicada y que requiere de mucha reflexión pues supone:

a) determinar las VI (su naturaleza y sus niveles) y VD (y su naturaleza) con las que vamos a trabajar.

b) Determinar cómo vamos a controlar las VE.

c) Determinar con cuántas muestras vamos a trabajar, de qué tipo (independientes, relacionadas, etc) y cuál va a ser el método de muestreo.

d) Determinar con qué técnica estadística analizaremos los datos (hay que saberla antes de llevar a cabo la medición). Es decir el tipo de análisis estadístico ha de estar claro antes de medir a los sujetos.

Como señaló Kerlinger (1991) el mejor diseño es aquel que satisface el “principio de MAX-MIN-CON”: a) maximiza la varianza tratamental (p.e. haciendo que las condiciones sean lo más distintas posibles => mayor tamaño del efecto); b) minimiza la varianza errática debida a diferencias individuales (p.e. dando correctamente las instrucciones,...) c) y controlar las variables extrañas (generalmente mediante la asignación aleatoria de los sujetos a las condiciones).

5) Determinación de con qué materiales (cuestionarios, tests, plantillas de registro observacional, etc) y aparatos (vídeo, ordenador, electroencefalogramas, etc.) trabajaremos, y cual será el procedimiento de medición (qué tareas deberán hacer los sujetos y su secuencialización).

- 6) Recogida de datos mediante mediciones fiables y válidas
- 7) Análisis estadístico de los datos recogidos (\*)
- 8) Interpretación de resultados: Discusión de los resultados en relación a las hipótesis planteadas y las teorías explicativas del problema.
- 9) Elaboración de un informe escrito (\*\*).

(\*) Principales pruebas estadísticas a aplicar para los distintos diseños de investigación:

A. Diseños observacionales: análisis de tablas de contingencia mediante  $\chi^2$ ; modelos log-lineal; fiabilidad inter-jueces (r, kappa, etc).

B. Diseños de encuestas o correlacionales: estadísticos descriptivos de todos los items (media, dt, ...); correlaciones entre los items (matriz de correlaciones); modelos de regresión para predicción; análisis factorial

C. Diseños experimentales y cuasi-experimentales:

**Nº de muestras a comparar:**

<b>VD:</b>	<b>1</b>	<b>2 ind</b>	<b>2 rel</b>
Cualitativa (nominal)	$\chi^2$ ; Kolmogorov	$\chi^2$ ; Kolmogorov	McNemar signos
Semi-cuantitativa (ordinal)	Wilcoxon	Mann-Whitney	Wilcoxon
<b>Cuantitativa (intervalo-razón)</b>	<b>t</b>	<b>t ind</b>	<b>t rel</b>

**Nº de muestras a comparar:**

<b>VD:</b>	<b>3 o más ind</b>	3 o más rel	2 o más VI (diseños AxB)
Cualitativa (nominal)	$\chi^2$ ; Kolmogorov	Cochran	$\chi^2$
Semi-cuantitativa	Kruskall-Wallis	Friedman	
<b>Cuantitativa (intervalo-razón)</b>	<b>ANOVA inter</b>	<b>ANOVA intra</b>	<b>ANOVA factorial intra, inter o mixto</b>

## \*\* 1.4. REDACCIÓN DEL INFORME

La elaboración del informe refleja secuencialmente los pasos seguidos a la hora de plantear el diseño una investigación. En concreto un informe psicológico debe contener (en este orden):

- **Título** (castellano e inglés)
- **Autores y filiación**
  
- **Resumen** (en castellano e inglés)
  
- **Introducción:** Planteamiento del problema. Revisión de la literatura sobre el mismo. Planteamiento explícito de las hipótesis.
  
- **Método:**
  - **Participantes**
  - **Materiales**
  - **Aparatos**
  - **Procedimiento**
  
- **Resultados:** Análisis estadísticos (descriptivos y/o inferenciales) de los resultados hallados. Tablas y gráficos, si proceden. ¿Los datos permiten aceptar o rechazar las hipótesis planteadas?
  
- **Discusión:** Implicaciones de los resultados para la hipótesis planteadas, y los modelos y teorías explicativos del problema de investigación.

- **Referencias:** Todo trabajo que en el texto del informe se cite alguna vez (p.e. Metlzoff, 2000) debe aparecer su referencia completa en el apartado de referencias . Hay tres tipos distintos de referencias:

- Libros, p.e.:

Pardo, A. y Sanmartín, R. (1995). *Análisis de datos en psicología II*. Madrid: Pirámide.

- Capítulos de un libro editado, p.e.:

Stinson, C.H. y Palmer, S.E. (1991). Pararell distributed processing models of person schemas and pathologies. En M.J. Horowitz (ed.), *Person schemas and maladaptive interpersonal patterns*. Chicago: University of Chicago Press.

- Artículos, p.e.:

Westmeyer, H. y Hagebock, J. (1992). Computer-assisted assesment: A normative perspective. *European Journal of Psychological Assessment*, 8, 1-16.

i) **Apéndices** (si proceden)



## 1.5. CRITERIOS PARA LA EVALUACION DE INFORMES PSICOLOGICOS

- 1) ¿El título del trabajo es preciso y claramente indicativo del trabajo de investigación que a continuación se presenta?
- 2) ¿El resumen describe adecuadamente lo que a continuación se presenta?. ¿Presenta conclusiones lícitas?

### EN LA INTRODUCCION:

- 3) ¿La pregunta o problema de la investigación está claramente formulada?
- 4) La revisión de la literatura ¿está actualizada?. ¿Demuestra el autor un conocimiento suficiente del problema de investigación?. ¿Se revisan adecuadamente las principales teorías y modelos existentes sobre el problema de investigación?
- 5) ¿Están clara y operativamente planteadas las hipótesis?. ¿Son relevantes en relación al problema de investigación?

### EN EL METODO:

- 6) ¿Se han seleccionado adecuadamente los participantes?. ¿La muestra es representativa?. ¿Es apropiado su tamaño?. ¿Son equivalentes los grupos en las VE?
- 7) ¿Los materiales y procedimientos están suficientemente detallados como para comprenderlos y poderlos replicar?.
- 8) ¿El diseño elegido es el indicado para someter a prueba las hipótesis planteadas?.
- 9) ¿Las variables independiente (y sus niveles) y dependiente son apropiadas?
- 10) ¿Hay un control adecuado de variables extrañas?

### EN LOS RESULTADOS

- 11) ¿Son apropiadas las pruebas estadísticas utilizadas?. ¿Hay errores en su presentación o en los cálculos?.
- 12) Las tablas y figuras ¿están descritas de forma clara?. ¿Son relevantes?. ¿Están bien hechas?. ¿Tienen sus respectivos encabezados o pies?.
- 13) ¿Los resultados están bien interpretados?

#### EN LA DISCUSION

- 14) ¿Son válidas las conclusiones?. ¿Están justificadas por los datos?. ¿Se discuten las implicaciones de los resultados hallados en relación a las teorías y modelos revisados en la introducción?.

#### EN LAS REFERENCIAS

- 15) ¿Están todas las referencias correctamente citadas?. ¿Falta alguna?.

#### OTROS ASPECTOS

- 16) ¿Hay consistencia entre el problema, las hipótesis, el diseño, los análisis y las conclusiones?
- 17) ¿Qué haría usted para mejorar la presente investigación?
- 18) Aspectos formales: señalar las erratas, palabras faltantes, etc. que haya encontrado en la lectura.

## 1.6. LA ETICA DE INVESTIGACION.

La investigación ha de respetar los derechos de los participantes (sean éstos personas o animales): No todo es lícito y justificable en nombre de la Ciencia. P.e. experimentos sobre obediencia de Milgram (1963; ver Davidoff, pp. 680-3; L & M, pp. 46-7), o grupos control sin tratamiento clínico.

El Colegio Oficial de Psicólogos de España ([www.cop.es](http://www.cop.es)) ha elaborado un código deontológico que todo licenciado en Psicología tiene la obligación de conocer y de cumplir (ver cuadro 2.13, pp. 48 de León y Montero):

- los sujetos han de estar bien informados de los objetivos de la investigación y de los riesgos que corren por participar.
- han de participar voluntariamente, dar su autorización (mejor por escrito) y poder abandonar cuando quieran si así lo desean
- evitar los procedimientos crueles y el sufrimiento físico y psicológico de los sujetos, siempre que sea posible
- evitar engañar a los sujetos y cuando esto no sea posible, decirles la verdad inmediatamente
- mantener la confidencialidad de los datos de los sujetos (secreto profesional)
- emplear tratamientos clínicos que han demostrado su eficacia experimentalmente (ver artículo de Labrador et al., 2003, [www.cop.es/vernumero.asp?id=1062](http://www.cop.es/vernumero.asp?id=1062))
- no falsificar datos (p.e. Burt), ni plagiar.

## TEMA 2. DISEÑOS EXPERIMENTALES.

Un experimento es una investigación: 1) donde el investigador tiene potestad para manipular directamente la(s) VI(s) (comparando al menos dos tratamientos: grupo control vs grupo experimental); y 2) donde la validez interna queda garantizada (generalmente formando grupos equivalentes mediante la asignación aleatoria de los sujetos a las condiciones experimentales). Recordemos que la validez interna se refiere a la seguridad en que la variación observada en la VD es debida exclusivamente a la manipulación de la(s) VI(s) y no al efecto de otras variables extrañas. La validez interna depende pues del grado de control de las variables extrañas (VE) o contaminantes. Para ello deberemos pues conocer las principales VE y cómo controlarlas (ver más abajo).

La validez interna es requisito imprescindible de todo experimento. Dicha validez interna nos permitirá garantizar la relación de causa-efecto entre la(s) VI y la VD.

En diseño experimental también son importantes:

- *Validez externa*: o posibilidad de generalización de nuestros resultados a otras poblaciones, otros contextos de investigación, otras condiciones, otros momentos temporales, etc. Es requisito aconsejable pero no imprescindible de un experimento. A veces la validez externa va en detrimento de la validez interna.
- *Validez estadística*: o validez de las conclusiones o inferencias derivadas de los análisis estadísticos. Depende entre otros factores del tamaño muestral ( $n$ ), del nivel de significación ( $1-\alpha$ ), de la potencia ( $1-\beta$ ) de la prueba estadística, del tamaño del efecto, etc.

La anterior definición de experimento excluye a los diseños cuasi-experimentales (manipulación por selección) y a los llamados diseños pre o pseudo-experimentales, que deberían usarse lo menos posible y sólo con fines exploratorios. P.e.:

- Diseño de grupo único con sólo medida post-tratamiento (p.e. evaluar el impacto de una terapia o campaña publicitaria).
- Diseño de grupo único con medidas pre y post-tratamiento
- Diseño de dos grupos no equivalentes con aplicación de un tratamiento a uno de ellos (p.e. analizar la eficacia de un método de enseñanza aplicándolo a los alumnos del colegio A y compararlo con los alumnos del colegio B que no reciben dicho método).

Todo experimento debe incluir como mínimo una condición control y una condición experimental.

## AMENAZAS A LA VALIDEZ INTERNA

Vamos a presentar ahora las principales VE que se involucran en cualquier investigación y cómo controlarlas (ver Meltzoff, cap.5):

### 1. Relativas a los participantes:

#### 1.1 Características personales y demográficas.

Se pueden controlar mediante:

- asignación aleatoria de muestras grandes (n>35)
- formar muestras homogéneas respecto a una o varias VE formando parejas, ternas, cuartetos, etc. (< validez externa)
- utilizar variables de bloqueo (=controlar VE desde el diseño formando muestras homogéneas)
- control estadístico a posteriori (ANCOVA)

#### 1.2. Efectos de *maduración e historia* de los participantes.

#### 1.3. *Motivación, “percepción del rol” y “actitud fastidiosa”* de los participantes.

#### 1.4. *Comunicación* entre los participantes

#### 1.5. Efecto de las *expectativas* de los sujetos (efecto placebo) => sujetos “ciegos”

#### 1.6. *Mortandad* experimental (aleatoria o no?)

#### 1.7. Fatiga

#### 1.8. Regresión a la media

### 2. Relativas al investigador:

#### 2.1. Variables demográficas (ensayos piloto)

2.2. Expectativas del experimentador (efecto Rosenthal o características de la demanda; p.e. Davidoff, pp 52) =>ciego y doble ciego.

### 3. Relativas al procedimiento, materiales y aparatos:

#### 3.1. Reactancia experimental (ensayos de práctica,...)

#### 3.2. Efecto de la práctica (p.e. pretest- post-test)

#### 3.3. Interacciones entre tratamientos múltiples

#### 3.4. Instrucciones ¿iguales y bien comprendidas?

#### 3.5. Efectos de suelo/techo

4. Relativas a jueces, observadores o expertos: entrenamiento adecuado (ensayos piloto), análisis de la fiabilidad interjueces ( $r > .80$ ).

5. Variables de ruido o causas inesperadas: preverlas y control estadístico.

Ya dijimos que una VI puede ser manipulada directamente o por selección. Otra forma de clasificar las VI consiste en distinguir entre:

- variables entre o intersujeto (*between subjects*), cuando utilizamos una muestra distinta de sujetos para cada uno de los niveles de la VI
- variables intrasujeto (*within subjects*), cuando utilizamos una única muestra de sujetos que recibe todos los niveles de la VI. Por eso también se las conoce como variables de medidas repetidas.

## TAXONOMIA DE LOS DISEÑOS EXPERIMENTALES

En base a la distinción entre VI inter o intra, y al número de VI manipuladas podemos clasificar los diseños experimentales así:

- a. 1 VI manipulada:
  - a1. inter ---> diseño unifactorial de X condiciones intersujeto o también llamado de X grupos al azar o completamente aleatorizados
  - a2. intra ---> diseño unifactorial de X condiciones intrasujeto o de X medidas repetidas
  
- b. 2 o más VI manipuladas:
  - b1. todas las VI inter ---> diseño factorial intersujeto o de grupos al azar AxB (p.e. 3x2, donde 3 y 2 indican el número de niveles de cada VI y el producto resultante -6- indica el número de condiciones o tratamientos experimentales distintos).
  - b2. todas las VI intra ---> diseño factorial intrasujeto o de medidas repetidas AxB.
  - b3. Una o varias VI inter y una o varias intra ---> diseño factorial mixto AxB.

En función del número de VD se suele distinguir también entre diseños univariantes (1 VD) y multivariantes (2 o más VD).

Cuando en un diseño inter bloqueamos (para controlarlas) una o varias variables extrañas se habla entonces de diseños de bloques al azar. Dicha variable de bloqueo se analiza luego como otra VI inter más.

Por último un tipo especial de experimentos son los experimentos n=1 o de caso único.

Tengamos presente que el uso de un diseño experimental no garantiza que estemos llevando a cabo un experimento: si nuestra investigación no tiene validez interna o la manipulación no es directa no estaremos haciendo experimentación. ---> El adjetivo “experimental” lo avala la validez interna, no el diseño utilizado.

## 2.1. Diseños intersujeto (o de grupos al azar o completamente aleatorizados).

Como su nombre indica exigen la asignación aleatoria de los sujetos a los tratamientos. Representan el concepto prototípico de experimento “clásico”. Dicha aleatorización garantiza la equivalencia de los grupos y por tanto el control de las VE, siempre que las muestras sean grandes ( $n > 30$  sujetos). Cuanto más grandes sean las muestras más garantías de control de las VE y mayor potencia tendrá la prueba estadística que luego utilizaremos.

Deberemos comparar siempre como mínimo un grupo experimental frente a un grupo control. Así cuando queramos comprobar si un tratamiento clínico (psicológico, farmacológico, etc.) es o no eficaz deberemos incluir siempre un grupo control que recibirá bien un placebo, bien el mejor tratamiento conocido a fecha de hoy, si se pueden plantear problemas éticos. De este modo podremos descartar la hipótesis de que la mejoría observada sea debida a las meras expectativas de sanación de los sujetos.

El problema que plantean los diseños intersujeto es que exigen muchos recursos materiales y humanos (pensemos p.e. que un diseño factorial  $3 \times 2$  ya exigiría 6 muestras de unos 30 sujetos = 180 sujetos como mínimo). Otro problema es al participar tantos sujetos la varianza de error aumenta, lo que minimiza la probabilidad de rechazar  $H_0$ . La solución a estos problemas es utilizar, siempre que se pueda, variables intrasujeto.

En los diseños factoriales hemos de analizar siempre si las VI interactúan o no:

### **El concepto de interacción.**

Se dice que dos variables interactúan (o que existe interacción entre ellas) cuando los efectos de los niveles de una de ellas afectan de modo distinto sobre cada uno de los niveles de la otra variable. Por contra dos variables no interactúan (o muestran efectos aditivos) cuando los efectos de los niveles de una de ellas se muestran constantes sobre los niveles de la otra

variable. Es decir la interacción AB representa el efecto conjunto sobre la VD de la aportación de A y B que no es reductible ni al efecto de A por sí solo ni al de B por sí solo. Por ejemplo si en medicina se dice que el hábito de fumar y el consumo de determinado fármaco *interactúan* sobre la salud, se quiere decir que la salud de los pacientes se ve afectada por el efecto *conjunto* de ambas variables, no teniendo por qué verse afectada por los efectos aislados de cada una de ellas.

Una forma intuitiva de entender el concepto de interacción es observar las gráficas relativas a las medias de los tratamientos. Imaginemos que estamos trabajando con un diseño factorial inter 3x2 y representáramos las medias de dichos 6 tratamientos, obteniendo alguno de los gráficos que aparecen en la figura 1.

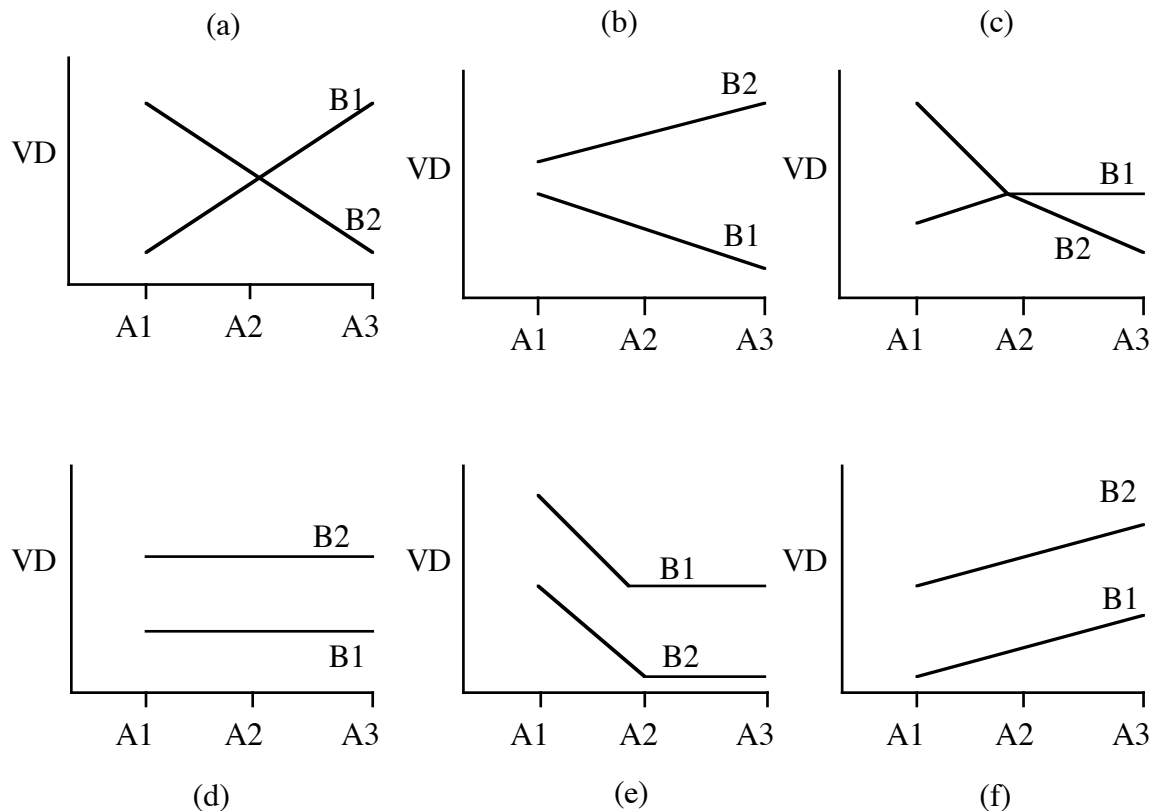


FIGURA 1.

En la figura 1 podemos observar cómo los casos a, b y c son casos prototípicos de interacción entre A y B, mientras que las figuras d, e y f son ejemplos claros de aditividad. En una gráfica *la ausencia de paralelismo en las líneas suele indicar interacción* entre variables, mientras que *el paralelismo suele denotar aditividad*. De todos modos la representación gráfica de las medias sólo nos da una idea intuitiva de si dos variables interactúan o no. Para confirmar dicha suposición deberemos aplicar el ANOVA.



Pueden existir así mismo interacciones de más de 2 variables. A las interacciones de dos variables se les denomina *interacciones de primer orden*. A las interacciones de tres variables se les denomina *interacciones de segundo orden*, y así sucesivamente. Las interacciones de segundo, tercer, cuarto,... orden son muy difíciles de interpretar. Es por ello que sólo se utilizan diseños de 1, 2 o 3 VI como máximo.

## 2.2. Diseños de bloques al azar.

Controlan una (o varias) variable extraña (VE) mediante la creación de grupos homogéneos respecto a ella. A dicha VE se la llama variable de bloqueo. El procedimiento a seguir para bloquear una VE es: 1) determinaremos los niveles a controlar de la VE (dos como mínimo), 2) formaremos grupos homogéneos de sujetos respecto a dichos niveles de la variable de bloqueo y 3) asignaremos aleatoriamente los sujetos de cada uno de los niveles de la variable de bloqueo a los niveles de la(s) VI. Dicha asignación aleatoria se encargará de controlar el resto de VE (p.e. si queremos comparar la eficacia de tres métodos de enseñanza podría ser conveniente bloquear la inteligencia de los sujetos en 2, 3, 4,...niveles de bloqueo).

La variable de bloqueo se analiza estadísticamente como si fuera una VI más (es decir analizaremos tanto sus efectos principales como si interactúa con la VI). Al ser los grupos más homogéneos la varianza de error del diseño disminuye, aumentando así la probabilidad de rechazo de la  $H_0$ . El principal problema de los diseños de bloques es la mortandad experimental que desequilibra los grupos.

Un caso especial de diseños de bloques al azar es cuando formamos parejas, ternas, cuartetos, ... de personas igualadas en alguna VE, y asignamos luego al azar los sujetos a los tratamientos. En este caso el número de niveles de la variable de bloqueo sería igual al número de parejas, ternas, cuartetos, ... con que contemos.

Hay otros tipos de diseños de bloques como los cuadrados latinos y cuadrados grecolatinos que no veremos aquí por su escaso empleo en experimentación.

### 2.3. Diseños intrasujeto o de medidas repetidas.

En ellos las VI son intrasujeto (una única muestra de sujetos pasa por todos los niveles de la VI), lo que conlleva una serie de ventajas: 1) más economía de recursos materiales y humanos involucrados en la investigación; 2) menor varianza de error lo que implica mayor probabilidad de rechazo de la  $H_0$ .

Entre los principales inconvenientes de este tipo de diseño están: 1) se inmiscuye siempre en ellos el efecto de la práctica que deberemos controlar por aleatorización intrasujeto o por contrabalanceo; 2) no todas las VI son susceptibles de ser manipuladas intrasujeto: p.e. las variables manipuladas por selección (cuasi-experimentación) son siempre entre; cuando los efectos de la VI son persistentes en el organismo de los sujetos (p.e. tratamientos farmacológicos o clínicos) dichas variables no pueden ser intra-sujetos. Es decir, sólo podrán ser VI intra aquellas que produzcan efectos en los sujetos que se eliminan completamente entre un tratamiento y otro.

### 2.4. Diseños factoriales mixtos.

Son aquellos en los que manipulamos al menos una VI intra y al menos una VI entre. Vale en ellos lo dicho para los diseños intra e intersujeto: analizar la interacción, controlar el efecto de la práctica, etc.

### 2.5 Diseños experimentales de caso único o $n=1$ .

Se utilizan en el ámbito de la psicología clínica (tradición del análisis experimental de la conducta), cuando no se puede trabajar con grupos, generalmente para verificar si determinada intervención clínica ha sido efectiva o no. No son tan aconsejables como los experimentos con grupos, pues su validez externa es limitada, pero si logramos establecer que las variaciones en la VD son debidas exclusivamente a la manipulación de la VI, desterrando otras causas, entonces tienen tanta validez interna (que no externa) como los experimentos convencionales. La validez externa se suele demostrar mediante distintas repeticiones de la investigación.

No hay que confundir los experimentos de caso único con los llamados estudios de casos, que no son experimentos (pues no hay control, ni validez

interna) si no investigaciones descriptivas de la evolución de pacientes clínicos propia corrientes clínicas psicoanalistas, humanistas, etc.

Entre las principales variables extrañas que pueden afectar a la validez interna de los experimentos de caso único caben destacar: maduración biológica de los sujetos (sobre todo cuando trabajamos con niños), historia personal del paciente, reactividad positiva del paciente ante el terapeuta, efecto del experimentador, mejoría por autosugestión, conductas patológicas cíclicas, etc.

Pero, repetimos, lo importante en este tipo de investigación es que el autor nos garantice la validez interna de la misma, es decir, garantías de control de las variables extrañas, pues si no hablaríamos de cuasi-experimentos  $n=1$ .

Estructura de los experimentos  $n=1$ . A pesar de que hay una gran variedad de diseños  $n=1$  (ver Barlow y Hersen, 1988 o Kazdin, 1988) todos ellos comparten tres fases básicas secuenciales:

a) Establecer una línea base de la conducta patológica. Para ello deberemos de tomar una serie de registros de la conducta patológica (VD) durante un periodo de tiempo suficientemente largo que nos permite analizar la tendencia y variabilidad de dicha conducta. Esta serie se suele llamar por convención fase A o línea base.

b) Aplicar la terapia que sea (refuerzos, castigo, relajación, fármacos, economía de fichas, tiempo fuera, aproximaciones sucesivas, etc) y contingentemente repetir la serie de registros sobre la misma conducta patológica (o fase B o tratamiento).

c) Interpretación de los resultados. Por tradición del análisis experimental de la conducta se suele hacer por medio de un simple análisis visual de las gráficas. De todos modos hay que cerciorarse de que tengamos suficientes registros por serie, prestando singular atención a los cambios de tendencia y de nivel de la variable dependiente, y desconfiando cuando ésta muestre mucha variabilidad. Hay pruebas estadísticas que nos pueden ayudar a salir de dudas, p.e. análisis de series temporales, técnica de división en mitades, etc. pero suelen ser difíciles de aplicar e interpretar.

Principales tipos de diseño  $n=1$ :

- AB o línea base-tratamiento. Es el más sencillo de aplicar pero tiene poca validez interna (por lo que es un diseño cuasi-experimental), dado que puede haber variables extrañas enmascaradas que expliquen la mejoría del paciente. Pero a veces es el único que se puede aplicar (por limitaciones temporales p.e.)
- BAB: se aplica cuando la intervención ha de ser inmediata (p.e. en conductas autodestructivas).
- ABA: Poco utilizado por que acaba en una fase sin tratamiento, lo que provoca el desinterés de los sujetos.
- ABAB o de "vuelta atrás". Es el más utilizado por ser el que más validez interna tiene, pues permite detectar recuperaciones espontáneas, por autosugestión, efecto del experimentador, etc.
- Diseños de línea base múltiple. Son una extensión de los diseños AB pero con distintas líneas base sobre distintas conductas (VD) a modificar (o también sobre la misma VD medida en distintos pacientes) a las que se aplica un mismo tratamiento. Su estructura general es del tipo
  - AB            sobre la 1ª conducta a modificar
  - AAB        sobre la 2ª conducta
  - AAAB      sobre la 3ª conducta
  - AAAAB, sobre la 4ª conducta , etc.

Por ejemplo a través de un diseño de línea base múltiple un maestro podría estar interesado en mejorar la capacidad lectora (VD1), capacidad numérica (VD2) y habilidades psicomotrices (VD3) de sus alumnos aplicando un sistema de economía de fichas según el esquema anterior. La ventaja de estos tipos de diseño es que tienen alta validez interna, pues observando las líneas base no tratadas podemos encontrar variables extrañas enmascaradas. En nuestro ejemplo si tras aplicar la economía de fichas sobre la capacidad lectora vemos que la línea base de la capacidad numérica aumenta querría decir que posiblemente hay una variable enmascarada detrás (p.e. la motivación).

Para aplicar este tipo de diseños hay previamente que verificar que 1) las distintas VD no esten correlacionadas, es decir, sean independientes y 2) sean igualmente sensibles a la misma VI o tratamiento.

# TEMA 3. ANALISIS DE DISEÑOS EXPERIMENTALES (Y CUASI- EXPERIMENTALES).

## 3.1. Conceptos previos: Nivel de significación, tamaño del efecto y potencia.

Los programas estadísticos nos dan junto al valor del estadístico calculado (F, t, z, ...) una p ( o sig.) que indica la probabilidad de rechazar  $H_0$  incorrectamente ( $\alpha$ ). Recordemos que dicha p debe de ser menor o igual a 0.05 para hablar de diferencias estadísticamente significativas.

Muchas veces se piensa incorrectamente que una p muy pequeña es indicativa de que la manipulación de la VI sobre la VD ha sido muy efectiva, es decir, de un tamaño del efecto muy alto. Y eso no es así pues p depende del tamaño muestral: una  $p=0.03$  podrá tender relevancia psicológica ante un  $n=30$  p.e., pero la misma p ante un  $n=3000$  no tiene ninguna relevancia.

La relación entre  $\alpha$ ,  $\beta$ , n y el tamaño del efecto se puede apreciar en la figura 1.

		Ho verdadera	Ho falsa
D E C I S I O N	Acepto $H_0$	Decisión correcta ( $1 - \alpha$ )	error tipo II ( $\beta$ )
	Rechazo $H_0$	error tipo I ( $\alpha$ )	Decisión correcta ( $1 - \beta$ )

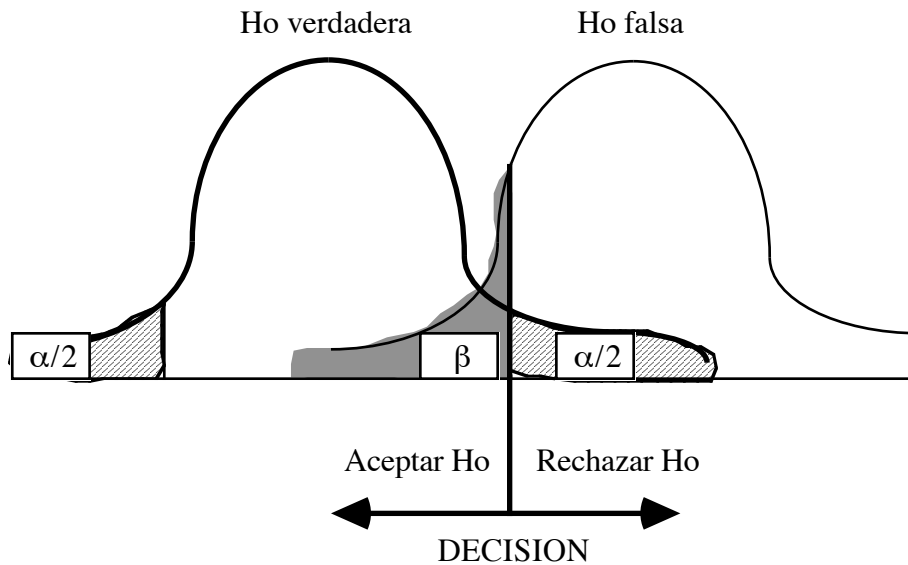


Figura 1.

La situación ideal es conseguir con muestras suficientes (p.e.  $n=30$ ) potencias y tamaños del efecto altos (p.e.  $>.80$ )

Con el programa SPSS podremos calcular la potencia de las pruebas ( $1-\beta$ ) y su tamaño del efecto ( $\eta^2$  -eta<sup>2</sup>- y  $\omega^2$  -omega<sup>2</sup>-).

### 3.2. El programa SPSS.

Recordemos que los datos se introducen en el editor de datos así: cada fila representa un sujeto distinto y cada columna una variable distinta. P.e. si comparamos dos muestras independientes de 12 sujetos cada una en una VD, utilizaremos 24 filas y 2 columnas (1 para la VD y otra para la VI para indicar el grupo: 1 vs 2). Pero si queremos comparar el rendimiento de 12 sujetos medidos 2 veces en una VD (o dos muestras distintas relacionadas o apareadas) utilizaremos 12 filas y 2 columnas (que representarán las dos mediciones en la VD).

Los comandos principales que utilizaremos en el SPSS son:

- a) Para hacer una prueba t para muestras independientes seleccionaremos el menú *analizar + comparar medias + prueba t para muestras independientes*.
- b) Para hacer una prueba t para muestras relacionadas seleccionaremos *analizar + comparar medias + prueba t para muestras relacionadas*.

- c) Para hacer un ANOVA unifactorial inter seleccionaremos *analizar + comparar medias + ANOVA de un factor*. Pediremos luego en *opciones* que nos haga test de homogeneidad (Levene) y en *post-hoc* las pruebas a posteriori que deseemos (p.e. Student-Newman-Keuls).
- d) Para hacer un ANOVA unifactorial intra seleccionaremos *analizar + modelo general lineal + medidas repetidas*. Luego en la ventana que nos salga deberemos nombrar el factor intra, sus niveles + *añadir + definir*.
- e) Para hacer un ANOVA factorial inter seleccionaremos *analizar + modelo general lineal + univariante*. Luego en la ventana pondremos en *dependientes* nuestra VD y en *factores fijos* nuestras VI. Pediremos luego en *opciones* que nos haga test de homogeneidad (Levene) y pruebas post-hoc sobre los efectos principales, así como que nos muestre las medias para las 2 VI así como para su interacción. Recordemos que el SPSS no hace pruebas de efectos simples por lo que si la interacción nos sale significativa deberemos hacerlas a mano trabajando con un  $\alpha = .05/\text{número de comparaciones}$  (Bonferroni).
- f) Para hacer un ANOVA factorial intra seleccionaremos *analizar + modelo general lineal + medidas repetidas*. Luego en la ventana que nos salga deberemos nombrar los factores intra, sus niveles + *añadir + definir*. En *opciones* pediremos que nos muestre las medias para las 2 VI así como para su interacción.
- g) Para hacer un ANOVA factorial mixto seleccionaremos *analizar + modelo general lineal + medidas repetidas*. Luego en la ventana que nos salga deberemos definir el o los factores intra, sus niveles + *añadir*. Luego pulsaremos *definir* y definiremos las VI inter y los niveles de las VI intra. En *opciones* pediremos que nos muestre las medias para las 2 VI así como para su interacción.
- h) Para aplicar las distintas pruebas no paramétricas seleccionaremos *analizar + pruebas no paramétricas* + elegiremos la apropiada para cada caso.

### 3.3. Análisis de diseños unifactoriales inter.

Vamos a asumir siempre que la VD es de naturaleza cuantitativa (escala de intervalo o de razón) o en su defecto semi-cuantitativa (ordinal).

3.3.1. Cuando la VI tiene dos niveles utilizaremos la prueba t para muestras independientes. Si la VD fuera semi-cuantitativa, o las muestras muy pequeñas utilizaríamos la prueba no paramétrica de Mann-Whitney.

En el SPSS para hacer una prueba t para muestras independientes seleccionaremos el menú *analizar + comparar medias + prueba t para muestras independientes*. Para aplicar las distintas pruebas no paramétricas seleccionaremos *analizar + pruebas no paramétricas + elegiremos la apropiada para cada caso*.

EJEMPLO 1: 20 personas son asignadas al azar a un grupo control y otras 20 a un grupo experimental. Introducimos los datos en el SPSS, aplicamos una *prueba t para muestras independientes* y obtenemos el siguiente output:

Variable	Number of Cases	Mean	SD	SE of Mean
GRUPO EXP	20	,6250	,988	,221
GRUPO CONT	20	-,0220	,748	,167

Mean Difference = ,6470

Levene's Test for Equality of Variances: F= 1,817 P= ,186

t-test for Equality of Means				95%	
Variances	t-value	df	2-Tail Sig	SE of Diff	CI for Diff
Equal	2,33	38	,025	,277	(,086; 1,208)
Unequal	2,33	35,40	,025	,277	(,085; 1,209)

Esto implica que hay diferencias significativas entre el grupo control y el experimental. Si el contraste fuera unilateral deberíamos dividir  $.025 / 2 = .0125$ .

Recordemos que los resultados se citan textualmente (y no añadiendo las tablas de salida del SPSS). En nuestro ejemplo diríamos algo así como “la diferencia entre el grupo experimental y el grupo control (medias de 0.63 y -0.02, respectivamente) fue estadísticamente significativa ( $t_{38}=2.33$ ,  $p=.025$ )”.

Si analizamos los datos anteriores con la prueba de Mann-Whitney obtendríamos el mismo resultado:



Mean Rank	Cases			
24,58	20	GRUPO = 1,00		
16,42	20	GRUPO = 2,00		
Exact		Corrected for ties		
U	W	2-Tailed P	Z	2-Tailed P
118,5	491,5	,0263	-2,2381	<b>,0252</b>

-----

3.3.2. Cuando la VI tiene tres o más niveles utilizaremos ANOVA (=análisis de la varianza) inter. Previamente deberemos analizar el supuesto de homoscedasticidad de los grupos mediante la prueba de Levene (que deberá darnos no significativa si hay homoscedasticidad). Si la F resultante ( $F = MC_{inter} / MC_{error}$ ) es significativa entonces deberemos aplicar una prueba a posteriori (p.e. Newman-Keuls, Tukey, Scheffé, LSD de Fisher, t de Student, Bonferroni, etc). Cuando la VD sea semi-cuantitativa, o el supuesto de homoscedasticidad no se cumpla es mejor utilizar el ANOVA no paramétrico de Kruskal-Wallis.

EJEMPLO 2. Asignamos al azar 5 sujetos a un grupo control (grupo 1), 5 sujetos a un primer grupo tratamental (grupo 2), y otros 5 a un segundo grupo tratamental (grupo 3). En el SPSS para hacer un ANOVA unifactorial inter seleccionaremos *analizar + comparar medias + ANOVA de un factor*. Pediremos luego en *opciones* que nos haga un test de homogeneidad de la varianza (test de Levene) y en *post-hoc* las pruebas a posteriori que deseemos (p.e. Student-Newman-Keuls, S-N-K). Los resultados fueron:

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	2	53,7333	26,8667	62,0000	<b>,0000</b>
Within Groups	12	5,2000	,4333		
Total	14	58,9333			

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
,6038	2	12	<b>,563</b>

(\*) Indicates significant differences which are shown in the lower triangle

Mean	GRUPO	
5,6000	Grp 1	
9,4000	Grp 2	*
9,8000	Grp 3	*

En primer lugar nos fijaremos si hay homogeneidad de varianzas en la prueba de Levene. Dado que  $.563 > 0.05$  quiere decir que dicho supuesto se cumple y por tanto podemos hacer el ANOVA. En dicho ANOVA vemos que la probabilidad asociada a la F calculada es menor que  $.05$ , luego podemos rechazar la  $H_0$  que afirma que las 3 medias tratamentales son similares. En otras palabras el ANOVA nos indica que la VI influye de forma significativa sobre la VD. Pero ahora deberemos analizar mediante la prueba posteriori de S-N-K entre qué medias existen diferencias. Si observamos los resultados veremos que hay diferencias entre los dos grupos experimentales con respecto al grupo control, pero no entre ellos.

Los resultados los citaríamos más o menos así: “ Un ANOVA inter determinó como significativo el efecto de la VI ( $F_{2,12}=62.00$ ,  $MSe=0.43$ ,  $p<.0001$ ). Una prueba a posteriori de Student-Newman-Keuls determinó como significativas las diferencias entre los dos grupos experimentales con respecto al grupo control, pero no las diferencias entre ellos”.

Si analizáramos los datos anteriores con una prueba no paramétrica (Kruskall-Wallis) llegaríamos a conclusiones similares.

### 3.4. Diseños factoriales inter.

Utilizaremos ANOVA factorial inter  $A \times B$ , previo análisis de la homoscedasticidad. Dicho ANOVA nos dará tres F, dos relativas a los efectos principales de A y B, y la otra relativa a la interacción  $A \times B$ . Si nos sale significativa alguna de F relativa a los efectos principales deberemos utilizar una prueba a posteriori. Si nos sale significativa la interacción deberemos aplicar una prueba de efectos simples (+ gráfica de la interacción). Dado que el SPSS no hace pruebas de efectos simples lo que haremos será analizar la interacción aplicando distintas pruebas t sobre las medias representadas en la gráfica trabajando con un  $\alpha = .05/\text{número de comparaciones}$  (Bonferroni).

EJEMPLO 3. Pensemos en un diseño donde manipulamos por un lado un grupo control y dos grupos tratamentales y por otro lado el sexo. Es decir estamos hablando de un diseño factorial inter  $3 \times 2$  donde asignamos al azar 5 sujetos a cada uno de los 6 tratamientos.

En el SPSS para hacer un ANOVA factorial inter seleccionaremos *analizar + modelo general lineal + univariante*. Luego en la ventana pondremos en *dependientes* nuestra VD y en *factores fijos* nuestras VI. Pediremos luego en *opciones* que nos haga test de homogeneidad y pruebas post-hoc sobre los efectos principales, así como que nos muestre las medias para las 2 VI así como para su interacción.

Las medias fueron:

	Grupo 1 (control)	Grupo 2 (experimental)	Grupo 3 (experimental)	TOTAL
mujeres	5.2	8.4	10.2	7.9
hombres	3.6	3	1.6	2.7
TOTAL	4.4	5.7	5.9	

Los resultados del SPSS fueron:

Univariate Homogeneity of Variance Tests

Cochrans C(4;6) = ,25000, P = 1,000 (approx.)

Bartlett-Box F(5;741) = ,35076, P = ,882

\*\*\*\*\* Analysis of Variance\*\*\*\*\*

Source of Variation	SS	DF	MS	F	Sig of F
WITHIN+RESIDUAL	27,20	24	1,13		
GRUPO	13,27	2	6,63	5,85	,009
SEX	202,80	1	202,80	178,94	,000
GRUPO BY SEX	61,40	2	30,70	27,09	,000
(Model)	277,47	5	55,49	48,96	,000
(Total)	304,67	29	10,51		

(\*) Indicates significant differences which are shown in the lower triangle

Mean		1	2	3
	GRUPO			
4,4	Grp 1			
5,7	Grp 2	*		
5,9	Grp 3	*	*	

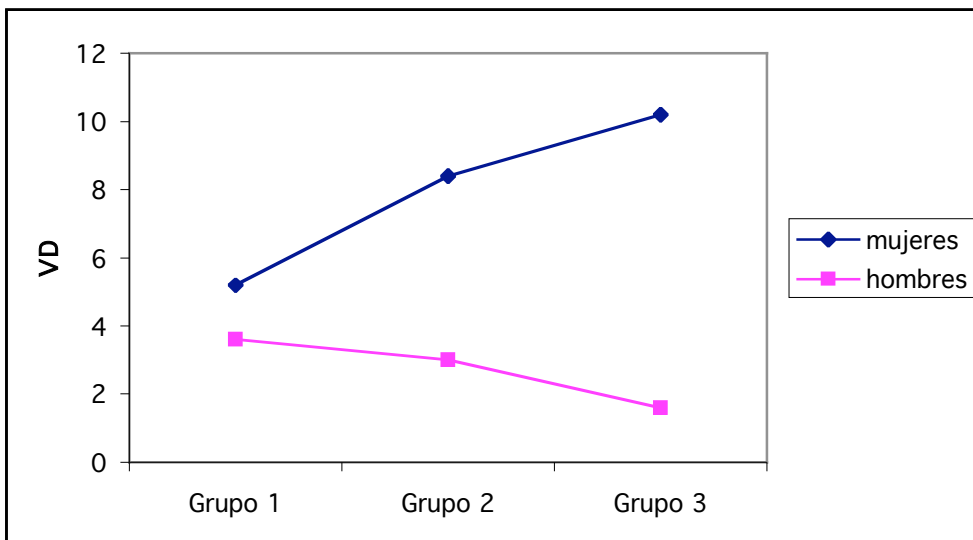
Source of Variation	Effect Size and Observed Power at the ,0500 Level	
	ETA Sqd	Power
GRUPO	,32784	,827
SEX	,88174	1,000
GRUPO BY SEX	,69300	1,000

En primer lugar se observa que hay homogeneidad de varianzas, luego procede el ANOVA.

En segundo lugar observamos que tanto la variable grupo, como la variable sexo, como la interacción de ambas son significativas, ( $F_{2,24}=5.85$ ,  $MSe=1.13$ ,  $p=0.009$ ;

$F_{1,24}=178.94$ ,  $MSe=1.13$ ,  $p<0.001$ ;  $F_{2,24}=27.09$ ,  $MSe=1.13$ ,  $p<0.001$ ), respectivamente. Debemos pues aplicar pruebas a posteriori:

- i. Sobre la variable sexo no procede aplicar una prueba a posteriori pues sólo hay 2 medias que comparar y su diferencia es significativa: las mujeres rinden más en la VD que los hombres (ver tabla de medias).
- ii. Sobre la variable grupo los resultados de la prueba a posteriori S-N-K nos muestran que hay diferencias entre el grupo control con los dos grupos tratamientos, no habiendo diferencias entre ellos.
- iii. Sobre la interacción sexo\*grupo lo mejor es hacer la gráfica (ver abajo) y 3 hacer pruebas t “a mano” comparando si hay diferencias entre hombres y mujeres en el grupo 1, en el grupo 2 y en el grupo 3 y tomando como significativas aquellas t con probabilidades iguales o menores de  $0.05/3 = 0.0167$ . Si las hacemos encontraremos que las probabilidades asociadas a dichas tres t son, respectivamente, .035, <.0001 y <.0001. Luego sólo existen diferencias entre hombres y mujeres en los dos grupos tratamientos.



Por último podemos observar en los análisis que la potencia y los tamaños del efecto paracen adecuados.

### 3.5. Diseños unifactoriales intra-sujetos.

- b. Cuando la VI tiene dos niveles utilizaremos la prueba t para muestras relacionadas. Si la VD fuera semi-cuantitativa, o las muestras muy pequeñas, utilizaríamos la prueba no

paramétrica de Wilcoxon. La interpretación de los resultados del SPSS es similar a la del ejemplo 1.

- c. Cuando la VI tiene tres o más niveles utilizaremos ANOVA intra (más pruebas a posteriori, si proceden). El equivalente al supuesto de homoscedasticidad en este tipo de diseños se llama supuesto de esfericidad (varianzas iguales y covarianzas nulas) y se analiza mediante el test de Mauchly (que deberá darnos no significativo si hay esfericidad). Cuando la VD sea semi-cuantitativa utilizaremos el ANOVA no paramétrico de Friedman. La interpretación de los resultados del SPSS es similar a la del ejemplo 2.

- Estas mismas pruebas se aplican en el caso de muestras distintas pero relacionadas o apareadas (por pares, tripletas, cuartetos, etc.).

### 3.6. Diseños factoriales intra.

Utilizaremos ANOVA factorial intra  $A \times B$ , analizando los efectos principales significativos con pruebas a posteriori, y los efectos de interacción significativos con pruebas de efectos simples. La interpretación de los resultados del SPSS es similar a la del ejemplo 3.

3.7. Diseños factoriales mixtos. Aplicaremos ANOVAS mixtos  $A \times B$  procediendo como anteriormente hemos descrito. Los supuestos de aplicación de este tipo de ANOVA se analizan aplicando el test de Box que analiza que se cumpla el supuesto de esfericidad en cada uno de los niveles inter. La interpretación de los resultados del SPSS es similar a la del ejemplo 3.

## 4. Cómo citar resultados:

Recordemos que los resultados se citan textualmente (y no añadiendo tablas de salida del SPSS) en el apartado del informe de “Resultados”. Ejemplos:

- “la diferencia entre hombres y mujeres fue estadísticamente significativa ( $t(31)=3.21, p=0.02$ )”
- “los efectos principales de la variable ‘tipo de estímulo’ fueron significativos ( $F(1,50)=4.56, \eta^2=0.52, p=0.001$ ), los efectos principales de la variable ‘frecuencia’ no fueron significativos, mientras que la interacción de ambas variables fue significativa ( $F(2,50)=3.56, \eta^2=0.62, p=0.03$ ; ver figura 1).

Cuando la  $p$  que acompaña a un estadístico está cercana a 0.05 (p.e. 0.07) se dice que es “marginamente significativa”, lo que quiere decir que ese estadístico sería significativo aumentando un poco el tamaño muestral.

Recientemente se está aconsejando incluir el tamaño del efecto (eta cuadrado:  $\eta^2$ ) junto al valor del estadístico citado. Dicho estadístico nos da un valor del tamaño del efecto en una escala de 0 a 1, no dependiente del tamaño muestral, cosa que sí ocurre en las  $p$  o sig. Tamaños del efecto de 0.80, 0.50 y 0.20 representan aproximadamente tamaños del efecto altos, medios y bajos, respectivamente. Además incluir el tamaño del efecto nos permite la posibilidad de realizar posteriormente meta-análisis. El **meta-análisis** es una técnica estadística descriptiva que permite calcular el tamaño del efecto promedio de un conjunto de publicaciones relativas a un mismo campo de investigación donde generalmente hay resultados contradictorios. P.e. Smith y Vela (2001) revisaron 93 experimentos relativos al llamado efecto del contexto (si cambia el contexto entre la fase de estudio y la fase de recuerdo o reconocimiento de unos materiales, ¿ello afecta a los resultados?) y encontraron que el tamaño del efecto promedio era de 0.27 que aunque moderado era estadísticamente significativo, lo que fue interpretado como que dicho efecto existe pero es difícil de elicitar. Para saber más sobre meta-análisis:

- Botella, J. y Gambara, H. (2002). *¿Qué es el meta-análisis?*. Madrid: Biblioteca Nueva.
- Botella, J. y Gambara, H. (2006). Doing and reporting a meta-analysis. *International Journal of Clinical and Health Psychology*, 6, 425-440.

## TEMA 4. INTRODUCCION A LA INVESTIGACION CAUSI-EXPERIMENTAL.

Se encuadran aquí dos tipos principales de investigación:

- a) aquella en que la VI se manipula por selección
- b) aquella en que la manipulación de la VI es directa pero no podemos asignar aleatoriamente los sujetos a los tratamientos (como ocurre en investigaciones de ámbito escolar, hospitalario, empresas, etc. donde no siempre es posible realizar experimentos).

Su objetivo es establecer relaciones causales entre VI y VD. Sin embargo con este tipo de diseños dichas relaciones de causalidad no pueden demostrarse al 100% (es decir no tienen tanta validez interna como los experimentos). En la medida en que el investigador sea capaz de ir descartando las hipótesis alternativas explicativas de la variación de la VD, en esa medida irá aumentando la validez interna (p.e. haciendo que los grupos sean lo más equivalentes posibles en todas las VE posibles, difiriendo sólo en la VI). Esta menor validez interna se suele compensar, generalmente (pero no siempre), con una mayor validez externa.

Se hace muy difícil sistematizar la amplia variedad de diseños cuasi-experimentales. Una posible taxonomía clasificatoria podría ser (Cook y Campbell, 1979):

A. Diseños pre-post tratamiento:

A1. *Diseño pre-post de un solo grupo*: ya dijimos que tiene tan poca validez interna que se le llama pre-experimental. Atención al efecto de la práctica!

A2. *Diseño pre-post con un grupo control no equivalente*, es decir, diseño mixto 2 (pre-post) x 2 grupos (control y experimental): Uno de los diseños cuasiexperimentales más utilizados. Para aumentar la validez interna debemos intentar que ambos grupos en el pre-tratamiento sean lo más equivalentes posibles el las VE. Si hay efecto tratamental, entonces en los resultados deberemos encontrar una interacción significativa, no habiendo diferencias en el pretest, pero sí en el post-test.

A3. Idem al anterior pero *con grupo control de cohorte (=generación) distinta*.

B. Diseños sólo post-tratamiento:

B1. *Diseño de grupo único*. Ya dijimos que tiene tan poca validez interna que se le llama pre-experimental, pero a veces sólo nos queda esta opción de diseño (p.e. en la valoración de programas en el ámbito de la psicología social: evaluar el impacto de una campaña publicitaria o una ley de educación).

B2. *Comparar dos grupos (o cohortes) en el post*.



C. Diseños de series temporales: Son diseños similares a los experimentos  $n=1$  pero trabajando con grupos: medimos la evolución de una muestra a lo largo del tiempo (=VI).

C1. *Diseño pre-post de un solo grupo*: medimos varias veces a una muestra antes del tratamiento y después del mismo para ver su evolución. Es pues similar al diseño visto en A1 pero con más mediciones de la VD.

C2. *Diseño pre-post con un grupo control no equivalente*. Es decir obtendremos dos series temporales: una AB (para el grupo experimental) y otra AA (para el grupo control). Para aumentar la validez interna debemos intentar que ambos grupos en el pre-tratamiento sean lo más equivalentes posibles en todas las posibles variables extrañas.

E. Diseños evolutivos: En Psicología evolutiva y del desarrollo se suelen emplear diseños en los que la VI es el tiempo o la edad. Todos ellos son diseños cuasi-experimentales pues la manipulación de dicha VI es por selección. Se suele hablar de diseños:

E1. Transversales: comparar varios grupos de edades distintas (o cohortes) en un momento temporal.

E2. Longitudinales: seguir la evolución en el tiempo de uno o varios grupos (diseños de series temporales).

## TEMA 5. INTRODUCCIÓN A LA METODOLOGIA DE ENCUESTAS.

(Este tema se puede ampliar matriculándose en la asignatura optativa *Encuestas y estudios de opinión*)

La metodología de encuestas (también conocida como metodología correlacional, psicométrica, selectiva o de tests y cuestionarios) tiene por objetivo describir variables (psicológicas, sociales, ...) y analizar sus relaciones mediante la aplicación de encuestas, tests o cuestionarios a muestras representativas de sujetos. Representa la “otra” forma de entender la Psicología científica (Crombach, 1957) siguiendo la tradición comenzada por Galton, Pearson, Spearman, Thurstone, Binet,... Se diferencia de las metodologías experimental y cuasi-experimental en que:

- no hay manipulación de variables, es decir no hay VI
- no hay control de variables extrañas
- la información se recolecta pasando a los sujetos tests, cuestionarios o encuestas, que son herramientas estandarizadas de recogida de información, compuestas por una serie de preguntas (o ítems), y que deben cumplir determinados requisitos psicométricos (fiabilidad y validez)
- no se intenta establecer relaciones de causa-efecto si no tan sólo describir y observar “qué va con qué” (relaciones)
- se hace hincapié en las diferencias individuales, es decir observar cómo los sujetos de la muestra difieren o se escalan en determinado rasgo
- las muestras suelen ser más grandes y representativas (utilizando para su selección métodos de muestreo)

### FASES EN LA ELABORACION DE UNA ENCUESTA:

Para elaborar una encuesta hay que decidir sobre una serie de aspectos importantes:

- qué es lo que queremos medir y cómo lo vamos a hacer (elaboración de preguntas, etc.), es decir diseñar la encuesta
- cuál es la población objeto estudio (y cómo seleccionar la muestra)
- cómo pasar la encuesta, es decir cómo organizar el trabajo de campo
- como tratar la información recogida (codificación y análisis estadístico) y qué conclusiones podemos sacar de los datos

Ampliaremos ahora cada uno de estos puntos.

- a) Diseño y elaboración de la encuesta: Supone una reflexión profunda sobre cada uno de estos puntos:
- Establecer claramente los objetivos de la encuesta (qué se quiere medir)
  - Operacionalizar dichos objetivos, lo que no siempre es fácil (p.e. cómo medir el desarrollo cognitivo). Esta fase es pues de fundamental importancia.
  - Elaborar preguntas concretas. Debemos considerar si las preguntas serán de alternativas cerradas (tipo cuestionario) o abiertas (tipo entrevista, las cuales complican la codificación y análisis posterior), etc. (ver León y Montero, pps. 117-122; Gambará pp 53).
  - Pasar este primer borrador a una muestra piloto para obtener información sobre fiabilidad, validez, análisis de ítems, preguntas sin responder o muy cargadas en las alternativas NS, NC, etc.
  - Revisión de la encuesta en base tantas veces como sea necesario hasta tener certeza de que todas las preguntas son necesarias y se entienden bien, su orden es el adecuado y que la encuesta es corta, fácil y motiva al sujeto a responder.
- b) Selección de la muestra. Debemos aquí:
- definir la población objeto de estudio prestando atención a cómo se estratifica dicha población (p.e. en aulas escolares, en municipios, etc) y qué se pretende comparar (p.e. hombre vs mujeres; municipios de más de cien mil habitantes vs por debajo de dicha cantidad, etc.)
  - definir cómo seleccionaremos la(s) muestra(s) para que sea representativa. La representatividad de las muestras se consigue por medio de técnicas de muestreo y un tamaño muestral adecuados.
  - por último no debemos olvidar que siempre que queramos describir las características específicas de una muestra concreta (p.e. ¿cuáles son los rasgos de personalidad de las adolescentes anoréxicas?) deberemos sopesar la necesidad de incluir un grupo control equivalente (p.e. de adolescentes no anoréxicas) para analizar si tales rasgos son idiosincráticos de aquel grupo o no.
- Tipos de muestreo:
- Probabilístico:
    - Aleatorio simple. El mejor pero sólo se emplea en poblaciones pequeñas.
    - Sistemático: supone coger al azar un sujeto comprendido entre 1 y N/n y luego replicar sumando N/n.

- Estratificado proporcional: calculamos la proporción de sujetos de cada estrato en relación a toda la población y luego multiplicamos dicha proporción por el tamaño de mi muestra  $n$  (concepto de *afijación*). Luego elegimos al azar dentro de cada estrato.
- Por conglomerados. Se aplica también en poblaciones estratificadas, pero a diferencia del muestreo estratificado (donde se muestrean todos los estratos), aquí no muestreamos todos los estratos si no sólo una muestra aleatoria de los mismos (es decir, los estratos que no entran en dicha muestra aleatoria no son encuestados).
- Polietápico: Combinación de todos los anteriores. P.e. siguiendo un procedimiento estratificado o por conglomerados afijar el tamaño muestral adecuado a cada estrato y luego elegir una ruta aleatoria por cuotas de edad, sexo, etc.
- No probabilístico (plantean problemas de validez):
  - Muestras accidentales (sesgadas. P.e. encuestas en periódicos)
  - Muestreo a propósito, p.e. para ensayos piloto.
  - Muestreo de poblaciones de difícil localización (p.e. drogodependientes, poblaciones clínicas, etc.)

#### - Tamaño muestral:

Depende entre otras cosas del tamaño de la población, del nivel de confianza, del error máximo o de muestreo dispuestos a asumir y  $P=Q$ . Con ello consultaremos cualquier tabla de los libros de estadística (ver p.e. León y Montero, 2003, pp. 111). P.e. para hacer una encuesta cuya población sea la ciudad de Valencia,  $1-\alpha=.95$ , y el  $EM= \pm 2 \rightarrow n=2500$

Tanto el tipo de muestro utilizado como los anteriores parámetros deben de aparecer claramente en la **ficha técnica** de la encuesta.

c) Organización del trabajo de campo.

Debemos decidir aquí cómo vamos a pasar a la encuesta a los sujetos: mediante un pase colectivo con instrucciones por escrito (tipo cuestionario), mediante entrevistas personales, por correo, por teléfono, por internet, etc. debiendo conocer las ventajas e inconvenientes de cada uno de estos medios (ver León y Montero, pps. 123-124).

d) Codificación y análisis estadístico de los datos recabados. Antes de analizar se recomienda:

- elaborar un libro de códigos que especifique los valores permitidos para cada variable y recodificaciones lícitas (ítems invertidos, etc).
- llevar a cabo un control de calidad de los datos analizando la coherencia de los mismos
- analizar si hay muchos valores faltantes (aleatorios o no?), su significado y cómo tratarlos (eliminación, imputación, etc.)

Una vez contrastada la calidad de los datos ya podemos llevar a cabo los análisis estadísticos. Las técnicas más utilizadas son análisis descriptivos de cada ítem, tablas de contingencia y análisis de correlaciones entre ellos (y en su caso aplicación de modelos de regresión) así como recodificar los ítems en nuevas variables que los resuman (llamados constructos o factores, generalmente aplicando Análisis Factorial).

No olvidar a la hora de interpretar los resultados que no podemos demostrar relaciones de causa y efecto (p.e. el que el 80% de los encuestados en una encuesta crea que el aumento de delincuencia sea debido a la inmigración no nos permite afirmar que la inmigración aumente –o sea la causa de- la delincuencia, tan sólo que hay una relación entre ambas variables).

No debemos olvidarnos de elaborar la **ficha técnica** de la encuesta donde se debe especificar, entre otras cosas, la población objeto de estudio, el tipo de muestreo empleado, el tamaño muestral, el nivel de significación con el que trabajamos y el error de muestreo, la fecha de realización, etc.

Dado su coste no se suelen hacer estudios de validación de las encuestas: la estrategia suele ser dar por válidos los datos mientras no se demuestre lo contrario. La fiabilidad por el contrario la podemos calcular con el SPSS.

## DISEÑOS DE INVESTIGACION DE ENCUESTAS.

La taxonomía más aceptada suele distinguir entre:

- a) diseños transversales: una muestra medida una sola vez. Son los más utilizados.
- b) diseños longitudinales, donde se persigue analizar cómo una población cambia en el tiempo (p.e. en su tendencia de voto). Cabe hablar de varios tipos:
  - b1. Cuando la misma muestra es encuestada más de una vez (p.e. antes y después de una campaña política o publicitaria) se habla de *diseños de panel*. Su principal problema es la mortandad.
  - b2. Más común es encuestar muestras distintas (pero equivalentes) en distintos momentos temporales (como se hace en los sondeos de intención de voto, barómetro del CIS). Se les llama *diseños de tendencias*.
  - b3. En los *diseños de cohorte* se estudia la evolución de una cohorte (o generación) a lo largo del tiempo y/o su contrastación con otras cohortes (p.e. educados en tiempos de Franco vs educados en la democracia).

## TEMA 6. INTRODUCCIÓN A LA METODOLOGIA OBSERVACIONAL.

La observación científica se distingue de la popular en que es:

- Sistemática y estructurada, es decir replicable por otro investigador
- fiable, es decir, que otro investigador en condiciones similares debería obtener registros similares
- válida, que mida sólo lo que realmente lo que pretendamos medir y no otra cosa
- y si es posible que no produzca reactividad en el sujeto observado y que las expectativas del investigador no afecten a los resultados (ambos aspectos afectarían a la validez).

La metodología observacional se caracteriza por:

- las conductas se observan en sus contextos naturales de aparición, con un objetivo descriptivo (y no demostrar efectos o causas)
- poca o nula intervención del investigador para observar conductas generadas espontáneamente
- se suele aplicar en poblaciones donde es difícil hacer (cuasi-) experimentos o encuestas como p.e. en población infantil, población con patologías severas (retraso mental, esquizofrenia, etc.) o en población adulta normal donde la presencia del investigador puede influir sobre los resultados. Es pues propia de áreas como la Psicología Evolutiva y del Desarrollo, Psicología Educativa, Psicopatología, Etología.
- aunque a veces las observaciones también pueden ser tomadas como VD en (cuasi) experimentos.
- se suele utilizar principalmente como método para sugerir hipótesis de investigación que luego serán puestas a prueba con (cuasi) experimentos o encuestas.

El investigador, antes de llevar a cabo una investigación observacional, ha de tener claro qué, cómo, cuándo y dónde y por qué (hipótesis) observar.

Para ello es muy útil documentarse sobre trabajos llevados a cabo sobre el mismo tema, y llevar a cabo un estudio previo asistemático, que nos ayudará a establecer las categorías de observación. Vayamos por partes:

A) QUÉ OBSERVAR:

En psicología podemos observar infinidad de cosas: animales, una sola persona (incluido uno mismo o autoobservación), las interacciones sociales entre parejas, familias o cualquier otro grupo social, rasgos psicológicos, conductas, etc.

Es imprescindible operativizar la conducta a observar (p.e. la conducta agresiva), es decir descomponerla en sub-conductas que la evidencien (p.e. pega al niño de lado, le empuja, le grita, le tira sus cosas al suelo, etc).

Hemos de tener en cuenta que el marco teórico desde el que planteemos la investigación influirá de forma determinante en la elección de las categorías de observación que eligamos. Es decir marcos teóricos distintos nos llevarán a crear categorías de observación distintas.

De igual modo nunca hay que observar si no tenemos clara qué pregunta o hipótesis queremos responder.

## B) CÓMO OBSERVAR:

- En primer lugar hemos de decidir sobre el grado de intervención del investigador: Podemos elegir entre no intervenir para nada en la observación (observación naturalista, que es la que provoca menos reactancia en los sujetos), intervenir lo menos posible (en la observación participante o en el estudio de casos clínicos) hasta intervenir plenamente (p.e. en experimentos de campo).
- En la actualidad es de muchísima utilidad el uso de video-cámaras y espejos unidireccionales para facilitar la observación sin intervenir en ella.
- Podemos elegir entre llevar a cabo un mero registro verbal de lo observado, pero tiene mayor fiabilidad y validez el utilizar **plantillas de registro (códigos de observación)** formadas por categorías excluyentes y exhaustivas y, si puede, situadas en un mismo nivel de análisis (ver consejos León y Montero, pps. 87-88; ejemplo Navas, pp. 448).

La documentación previa y realización de uno o varios registros previos asistemáticos nos darán retroalimentación sobre la adecuación o no de las categorías que hemos elegido.

- Deberemos decidir así mismo si vamos a medir:
  - Ocurrencias (sí-no)



- Frecuencias (absolutas o relativas, éstas permiten la comparabilidad entre estudios).
- Latencias (tiempo que tarda un sujeto en responder a un estímulo)
- Duración de una conducta
- Intensidad de una conducta

### C) CUANDO OBSERVAR:

Podemos emplear diseños longitudinales o transversales.

Lo ideal sería observar siempre al objeto de estudio pero eso casi nunca es posible. Por ello se suelen utilizar muestreos de tiempo, bien sistemáticos o fijos (p.e. siempre a la misma hora o coincidiendo con determinadas conductas), bien aleatorios (empleando métodos de muestreo). Y pudiendo elegir también entre muestreos continuos (registramos todo lo que ocurre en un determinado periodo de tiempo) o por intervalos (intercalando periodos de observación con periodos de registro).

### D) DONDE OBSERVAR:

Cabe aquí lo dicho arriba. Casi nunca es posible observar en todas las situaciones al sujeto observado, por lo que se suele recurrir a los muestreos de situaciones, bien sistemáticos, bien aleatorios.

**ANALISIS DE LA FIABILIDAD.** Es imprescindible que nuestras observaciones sean fiables. Para comprobarlo lo que se hace es analizar el grado de acuerdo o consenso entre dos o más observadores (entrenados debidamente y preferiblemente “ciegos” para controlar el efecto Rosenthal o de las expectativas): Si la variable observada es cuantitativa o semi-cuantitativa se utiliza el coeficiente de correlación de Pearson, mientras que si es cualitativa se utiliza el coeficiente Kappa de Cohen (León y Montero, pp. 96), que tiene en cuenta la probabilidad de coincidir por puro azar. En cualquier caso se exige que la fiabilidad interjueces sea como mínimo de 0.80.

La validez de la observación depende sobretodo de la calidad del muestreo de situaciones y momentos elegido. Dos fuentes de error que afectan a la validez son:

- a) la reactividad en el sujeto observado (que se puede reducir habitándolo a la presencia del investigador, utilizando aparatos de grabación, espejos unidireccionales, etc.)
- b) las expectativas del investigador que pueden sesgar los registros. La mejor manera de controlar esta variable es trabajar con dos o más investigadores “ciegos”.

Un diseño observacional se puede analizar empleando infinidad de técnicas estadísticas: estadísticos descriptivos, tablas de contingencia,  $r$ , modelos log-lineal, etc.

## METODOLOGIAS CUALITATIVAS.

Se encuadran aquí un heterogéneo grupo de métodos de investigación descriptivos cuyo denominador común es el rechazo del análisis cuantitativo de datos y de la estadística, priorizando estrategias de investigación no estructuradas y subjetivas. Debido a la falta de objetividad que conllevan para muchos autores no merecerían el nombre de “métodos” sino que serían un “cajón de sastre” donde incluir aquellas estrategias de investigación no encuadrables en la metodología cuantitativa. Su mayor facilidad de empleo (apenas requieren preparación de materiales, preguntas, instrumentos, diseños, no hay que hacer análisis estadísticos, etc.) no debe de ser la justificación para no emplear cuando sea posible técnicas cuantitativas, pues la fiabilidad/objetividad de éstas siempre es mayor. Sirven muy bien, eso sí, para generar hipótesis analizables después con métodos cuantitativos.

Debido a su heterogeneidad se hace muy difícil hacer una taxonomía de estos métodos, pero destacaríamos:

a) La entrevista (en profundidad, abierta o no estructurada o en general el estudio de casos) pretende llegar a conocer “hasta el fondo” la idiosincrasia de un individuo o paciente (pues se emplea mucho en el ámbito clínico). Para ello el psicólogo debe de generar un clima de empatía que facilite la comunicación mutua. En un número indeterminado de sesiones (de unos 45 minutos) el psicólogo, mediante preguntas abiertas, deberá ir profundizando en lo que le interesa, comenzando por preguntas generales hasta las más específicas. Las sesiones acabarán cuando el psicólogo se dé cuenta de que la información obtenida es ya redundante. La experiencia y capacidad de empatía del psicólogo son puntos básicos.

b) El grupo de discusión es una entrevista abierta a un grupo de entre 5 y 10 personas (que no se conozcan entre sí y que intente ser una muestra representativa de la población objeto de estudio) en sesiones de 90 a 120 minutos. Se emplea mucho en Psicología Social (investigación de mercados, publicidad,...) para analizar los roles, actitudes y motivaciones de los distintos individuos. El psicólogo deberá plantear y enmarcar un tema de discusión y generar la confianza para que los individuos vayan mostrando sus opiniones divergentes. Su labor será motivar a los individuos más tímidos, frenar a los que monopolizan la discusión, redirigir la discusión cuando alguien se salga del tema, mostrar contradicciones, hasta llegar al final a una serie de conclusiones resumen de lo acaecido. Es importante acondicionar bien el lugar de la reunión (utilizar mesas redondas, sistemas de grabación...) haciéndolo lo más cómodo posible.

c) Método Delphi. Consiste en obtener la información más consensuada posible de un grupo de expertos (entre 10 y 30, separados geográficamente generalmente) en relación a un objetivo. Se emplea mucho en evaluación de programas, toma de decisiones, etc. La secuencia consta como mínimo de tres cuestionarios. El primero plantea preguntas abiertas sobre el tema de estudio con el objetivo de generar la mayor diversidad posible de respuestas. Con ellas se generan afirmaciones sobre las que se pide su grado de acuerdo en un segundo cuestionario. Un tercer cuestionario sintetiza los resultados anteriores y los hace conocer a todo el grupo de expertos con el fin de que hagan un esfuerzo por acercarse a las posiciones de la mayoría. El proceso se repite hasta llegar a un consenso razonable.

d) En la investigación etnográfica el investigador quiere conocer a fondo un determinado grupo social o cultural homogéneo (p.e. determinada minoría étnica, una “tribu urbana”, etc.) introduciéndose y conviviendo con ellos mediante la observación participante y la entrevista abierta.

e) La investigación-acción tiene sus orígenes en la Psicología Social de la Gestalt de Lewin y en la Psicología Social Comunitaria. Su objetivo es analizar un grupo social (generalmente con penurias o deficiencias, p.e. grupos de riesgo o grupos marginales) con el fin de transformarlo y mejorarlo mediante la aplicación de una serie de pasos que se retroalimentan entre sí: conocimiento de la comunidad, elaboración de un plan de acción, implantación y evaluación.