

εσταδιστιχ̄

εσταδιστιχ̄

# Psicometría

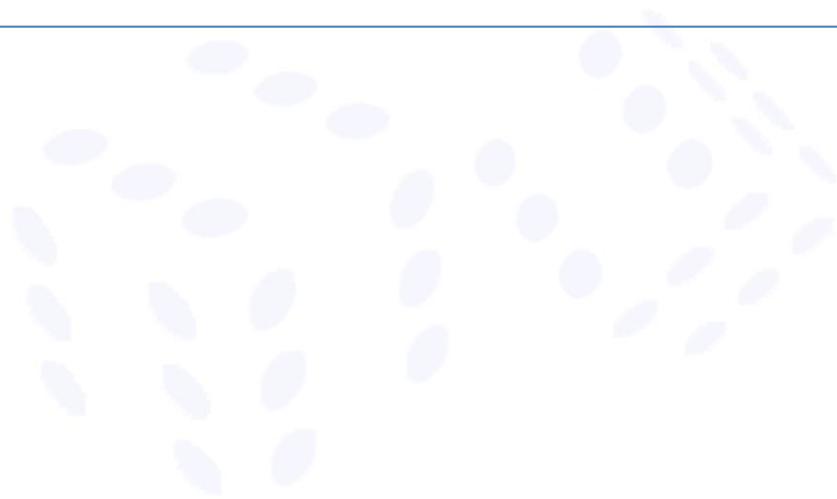
---

*Grado en Psicología UAM*

# ΕΣΤΑΔΙΣΤΙΧ

*Apuntes*

---



## 1. TEORÍAS DE LA MEDICIÓN

### DEFINICIÓN DE PSICOMETRÍA

La **Psicometría** es la disciplina que tiene por objeto la **medición de la "psique"**, de las capacidades, atributos o características psicológicas. Su objetivo es establecer teorías y métodos para la medición en Psicología y en otras disciplinas afines. Es objeto de la Psicometría el análisis de cualquier **asignación numérica que se realiza sobre variables psicológicas**, obtenida por diversos procedimientos o instrumentos. **Los test son los instrumentos más utilizados** para la obtención de puntuaciones. Sin medida no hay diagnóstico preciso, y sin diagnóstico preciso no hay intervención eficaz.

### TEORÍAS DE LA MEDICIÓN

Estudian los requisitos lógicos y matemáticos necesarios para poder medir y estudiaremos tres:

- Concepción conservadora de Campbell:** medir es asignar números a objetos, de modo que indiquen las cantidades de la propiedad del objeto medido. Bajo esta perspectiva la medición de lo psicológico es imposible.
- Teoría de las escalas de medición de Stevens:** medir es asignar números en base a una regla (ej: que la temperatura sea 0 no quiere decir que no haya temperatura). Los números no indican cantidades, indican relaciones. Define los niveles de medida nominal, ordinal, de intervalo y de razón.
- Teoría representacional de Suppes y Zinnes:** estudia las condiciones que debe satisfacer el sistema relacional empírico.

### ESCALAMIENTO

En el ámbito de la psicofísica surge el **primer problema** con la **medición de las sensaciones**, dentro del objetivo general de la **psicofísica** establecido por **Weber y Fechner**, que establecen la ley logarítmica para describir la relación matemática entre un continuo estimular,  $E$ , y la sensación producida en la persona,  $S$ . En el laboratorio de Leipzig, fundado por Wundt, algunos investigadores como **Donders y Cattell** empleaban el tiempo de reacción y la cantidad de eventos recordados como medida de rendimiento en diversos procesos psicológicos (atención, memoria...).

### MEDICIÓN DE LAS DIFERENCIAS INDIVIDUALES

**Galton** funda el laboratorio antropométrico de Londres, donde estudia las diferencias individuales, tratando de estudiar las **distribuciones de diversas medidas físicas y psicológicas de las personas**. Empleó instrumentos para medir la fuerza física o la capacidad de discriminación auditiva y visual. También le interesó medir los errores que cometían las personas en ciertas tareas de estimación de longitudes o ángulos. Uno de sus colaboradores, **J. M. Cattell**, en 1890 usó el término **"test mental"** para referirse al rendimiento en diversas tareas sensoriomotoras en las que se cuantificaba el tiempo de reacción o las estimaciones subjetivas.

### MEDICIÓN DE LAS CAPACIDADES COGNITIVAS SUPERIORES

**Binet** establece procedimientos para diferenciar a los niños con deficiencia mental (que no son capaces de conseguir los objetivos escolares) de los que no tienen problemas. Binet y Simon propondrán tests de inteligencia para medir capacidades cognitivas mediante el baremo cronológico de **cociente intelectual (CI)**. También en este contexto comienzan algunos esbozos de lo que serán los tests psicométricos de personalidad, cuyo arranque será más tardío debido en parte a que inicialmente la medición de la personalidad se realizaba con tests proyectivos.

## ANÁLISIS FACTORIAL

Es una técnica multivariante (que trabaja con muchas variables) de análisis de datos que permite pasar de un conjunto mayor de variables observadas a uno menor de variables latentes. Es una técnica de reducción de la dimensionalidad, y nos permite estudiar la validez de los constructos que medimos.

## TEORÍAS DE LOS TEST: TEORÍA CLÁSICA DE LOS TEST (TCT)

El padre de la Teoría clásica de los test es **Spearman** en 1904. Propone un sencillo modelo matemático:

$$X = V + e$$

Establece procedimientos para estimar la fiabilidad de las puntuaciones (bajo la definición de test paralelos) y las relaciones entre fiabilidad y longitud del test. La TCT, que será completada progresivamente por las aportaciones de otros autores, entre los que cabe destacar a **Thurstone y Cronbach**.

## TEORÍAS DE LOS TEST: TEORÍA DE LA RESPUESTA AL ÍTEM (TRI)

Se fundamenta en el **estudio de los ítems**: cómo se comporta cada ítem (no tanto del test), midiendo las propiedades métricas de los ítems (sin que dependan de las características particulares de las muestras) y de las estimaciones de los niveles de las personas (sin que dependan de las características particulares de los ítems aplicados). El modelo matemático es muy complejo. Es una teoría **más reciente y más completa** que supera alguna de las limitaciones de la TCT.

## TIPOS DE TEST

Consideramos genéricamente **un test** como un procedimiento estandarizado que sirve para muestrear ciertas conductas de las personas a partir de sus respuestas a una serie de indicadores (ítems) e inferir su nivel verdadero en el constructo o variable latente.

Para hablar propiamente de test, se ha de cumplir:

- Los ítems han de tener **propiedades psicométricas conocidas** y se han elaborado siguiendo las normas.
- Requiere una **administración controlada** siguiendo las instrucciones, las condiciones de administración y de puntuación establecidas en el proceso de construcción del test.
- Debe **pretender evaluar a una persona en un atributo** de interés. A veces se confunden los tests con otros instrumentos, como las encuestas, cuya finalidad no es puntuar a cada persona que responde, sino saber, por ejemplo, las tendencias de voto a los distintos partidos.

• Según....

✓ **El tipo de rendimiento: óptimo/ típico.**

✓ El medio de aplicación: lápiz y papel/ informatizados.

✓ El modo de aplicación: individual/en grupo

✓ El modo de interpretar las puntuaciones: normativos /referidos a criterio.

✓ El formato de respuesta: seleccionada/construida/ opción binaria/ opción múltiple/ categorías ordenadas/ipsativos

✓ El establecimiento de tiempo límite: velocidad/potencia.

✓ El marco teórico: proyectivos, comportamentales, situacionales...

## TIPO DE RENDIMIENTO

-**Óptimo**: el examinado intenta conseguir la puntuación más alta posible. Son test de conocimiento, de aptitudes. Hay respuestas correctas e incorrectas (Ej. Test de inteligencia, exámenes, etc.)

-**Típico**: el examinado intenta que sus respuestas reflejen cómo es él. No pretende obtener la máxima puntuación. No existen respuestas correctas o incorrectas. Son test de actitudes.

## MODO DE INTERPRETAR LAS PUNTUACIONES

- **Puntuación normativa (relativo):** compara al examinado con respecto a su grupo de referencia.

Ejemplo: test de inteligencia entre niños.

- **Puntuación referida a criterio (absoluto):** me indica qué cosas es capaz de hacer el examinado.

Ejemplo: Exámenes de inglés B2, C1... (obtengo el b2 cuando soy capaz de cumplir los criterios del examen).

## FORMATO DE RESPUESTA

-**Seleccionada u opción múltiple:** se tiene que elegir una alternativa entre las varias (rendimiento óptimo).

Es como un examen tipo test. Aunque sea de opción múltiple la corrección normalmente es dicotómica (0-1)

-**Construida:** se tiene que generar la respuesta que puede ser completar una frase, realizar un cálculo, etc.

-**Opción binaria:** se elige entre dos opciones, como verdadero/falso o si/no.

-**Respuesta graduada o de categorías ordenadas:** escala Likert (grado de acuerdo con el enunciado).

-**Ipsativos o elección forzosa:** tienes que seleccionar la opción que más te represente. La clave es que las opciones se formulan intentando controlar la deseabilidad social.

## TIEMPO LÍMITE

-**Test de velocidad:** los ítems son muy parecidos en dificultad, por lo que las puntuaciones se basan en cuántos ítems se han podido responder bien en el tiempo asignado.

-**Test de potencia:** la dificultad de los ítems va cambiando y observamos quién ha sido capaz de realizar los ítems más complejos. Se da un tiempo apropiado para que la mayoría de examinados pueda terminarlo.

## MARCO TEÓRICO DE REFERENCIA

-**Test conductual:** el ítem consiste en una tarea/problema que el examinado debe resolver.

-**Test situacional:** se plantea una situación y se presentan varias respuestas. Hay opciones correctas, parcialmente correctas y erróneas ya que es un test de rendimiento óptimo.

## 2. CONSTRUCCIÓN DE TESTS Y ANÁLISIS DE ITEMS

### 2.1 CONSTRUCCIÓN DEL TEST

Para construir un test hay que definir el **constructo a medir**, redactar los ítems (más de los que se utilizarán), pasarlo a una muestra y hacer el análisis de ítems, de fiabilidad, validez, e interpretación de puntuaciones.

Hay que decidir qué población nos interesa y cómo recogeremos la muestra, el contexto de aplicación, los objetivos del test, las adaptaciones en la aplicación del test, los medios, la longitud óptima del test...

En la **tabla de especificaciones** se detallan los contenidos del test, las destrezas cognitivas a evaluar y el porcentaje de ítems que debe corresponder a cada celdilla que resulta de cruzar los contenidos y las destrezas.

### TIPOS Y FORMATOS DE ITEMS

Los test de **rendimiento típico (RT)** y de **rendimiento óptimo (RO)** difieren en tres características:

**-El tiempo de administración del test:** en los test de RO, hay que determinarlo con mucho cuidado y es diferente también si el test es de velocidad o de potencia.

**-Tratamiento de las preguntas dejadas sin contestar:** en un test de RO, cuando no se responde a un ítem, puede ser que no se ha tenido tiempo de leer (ítem no alcanzado, test de velocidad) o que no haya querido dar la respuesta tras haberlo leído (omisión, test de potencia). En los test de RT las no respuestas suelen indicar que el ítem no se entiende o hay desinterés.

**-Respuestas al azar y sesgos de respuesta:** en los tests de RO con ítems de opción múltiple es posible acertar al azar mientras que en los tests de RT no deberían haber respuestas al azar pero podemos encontrar sesgos de respuesta, como la tendencia a utilizar las categorías extremas, la aquiescencia y la deseabilidad.

En los tests de **RO** los ítems pueden ser **preguntas abiertas** o **preguntas con opciones preestablecidas** como los ítems con respuesta seleccionada son los ítems de verdadero-falso y los de opción múltiple.

En los tests de **RT** los ítems pueden ser **de opción binaria** cuando la persona debe elegir entre dos opciones antagónicas como un V-F o de **categorías ordenadas**. Hay varios tipos de escalas de respuestas, como las tipo Likert donde se ha de manifestar el grado de acuerdo con la frase (de muy de acuerdo a muy en desacuerdo), y aquellas en las que se ha de indicar la frecuencia del comportamiento descrito en el enunciado. En las escalas Likert se suele recomendar que el número de categorías sea 5 o un valor próximo, que las categorías estén equilibradas y que el salto entre cada dos etiquetas consecutivas sea de similar cuantía.

### REDACCIÓN DE ITEMS DE OPCIÓN MÚLTIPLE

**- Elección del contenido que se desea evaluar:** cada ítem debe evaluar el contenido de una celdilla de la tabla de especificaciones, lo que garantiza que el test muestreará bien todo el contenido a evaluar. El ítem puede ser sencillo o complejo, concreto o abstracto, memorístico o de razonamiento.

**- Expresión del contenido en el ítem:** lo central debe expresarse en el enunciado y cada opción es un complemento que debe concordar gramaticalmente con el enunciado. La sintaxis o estructura gramatical debe estar ajustada al contenido y a la comprensión lingüística de las personas.

**- Construcción de las opciones:** la opción correcta debe ser sólo una, estar en distintas ubicaciones e ir acompañada de distractores plausibles. Las opciones deben ser preferiblemente tres, presentarse en vertical, ser autónomas entre sí. Debe evitarse las opciones "Todas las anteriores" y "Ninguna de las anteriores".

## REDACCIÓN DE ITEMS DE CATEGORÍAS ORDENADAS

**Recomendaciones** que pueden ayudar a la redacción en los test de rendimiento típico:

- Utilizar el tiempo presente.
- Redactar frases que darían respuestas distintas a los que tengan alto y bajo nivel en el rasgo.
- Que el contenido sea claro y concreto: frases cortas, simples e inteligibles.
- Que hayan ítems de modo directo e inverso para evitar la aquiescencia.
- Evitar el uso de negaciones, y de universales (todo, siempre, nunca...), pues llevan a elegir la misma categoría. Algunas "Escala de sinceridad" utilizan estos universales para detectar falseamiento de respuestas.

## INCLUSIÓN DE CATEGORÍA CENTRAL

En las escalas que miden frecuencia, la categoría central está recomendada. La opción de la categoría central se usa más cuando las actitudes sobre el tema son poco intensas o si se tiene poca información sobre el tema, pero puede servir como refugio en opiniones conflictivas.

## CUANTIFICACIÓN DE RESPUESTAS

Consiste en asignar un número a la respuesta a cada ítem, de forma que a mayor nivel de rasgo más nivel en el ítem. El número debe tener una relación con lo que mido.

**-Cuantificación dicotómica:** opción múltiple, verdadero/falso, si/no. Las respuestas correctas se cuantifican con un 1 y las incorrectas con un 0. La puntuación total en el test sería la suma de aciertos y fallos.

**-Cuantificación politómica:** en lugar de tomar solo dos valores las opciones pueden tomar más. Tener en cuenta el número de categorías en esta cuantificación. Además de con qué tipo de ítem estoy trabajando. Cuando cuantifico tengo que tener en cuenta si el ítem es directo o indirecto.

## 2.2 ANÁLISIS DE ITEMS

### MEDIA Y VARIANZA

En un test de RT la **media** debe estar entre 1 y k mientras que en los de RO estarían entre 0 y 1. La **varianza máxima** se daría si la mitad eligieran la opción más baja y si la otra mitad la opción más alta:

$$S_{max}^2 = 0,5 \cdot (x_{min} - \bar{x})^2 + 0,5 \cdot (x_{max} - \bar{x})^2$$

*Ejemplo: en un ítem de 4 opciones ¿cuál sería la varianza máxima?*

### ÍNDICE DE DIFICULTAD O PROPORCIÓN DE ACIERTOS (ID o p)

Un ítem es fácil o difícil en función del número de personas que acierten o fallen con respecto al total de respuestas. El índice de dificultad es la proporción de aciertos respecto al total de respuestas:

$$p_j = \frac{A}{N} \quad \bar{x} = \sum p_j = \frac{\sum x_i}{N} \quad S_i^2 = p_j \cdot (1 - p_j)$$

El ID va de 0 a 1 donde 1 es lo más fácil posible (todo el mundo acierta) y 0 es lo más difícil. Es un índice de "facilidad". Al diseñar un test de RO, se suele recomendar que se sitúen al inicio algunos ítems fáciles, en la parte central los de dificultad media (valores p entre 0,30 y 0,70) y al final, los más difíciles.

El índice de dificultad corregido por azar tiene en cuenta la posibilidad de que los sujetos acierten por azar:

$$p_j^c = \frac{P_c}{N} = \frac{A - \frac{E}{k-1}}{N}$$

donde  $\left\{ \begin{array}{l} k \rightarrow n^{\circ} \text{ de alternativas} \\ E \rightarrow n^{\circ} \text{ errores} \\ N \rightarrow n^{\circ} \text{ de sujetos} \\ P_c \rightarrow \text{Puntuación corregida} \end{array} \right.$

Ejemplo: con la siguiente tabla calcula:

- El índice de dificultad y la varianza del ítem 1, 2 y 3
- La puntuación media del test
- La dificultad corregida por azar del ítem 6.

|   | Ítems |   |   |   |   |   |
|---|-------|---|---|---|---|---|
|   | 1     | 2 | 3 | 4 | 5 | 6 |
| 1 | 0     | 1 | 0 | 1 | 0 | 1 |
| 2 | 1     | 1 | 1 | 0 | 0 | 1 |
| 3 | 0     | 1 | 0 | 1 | 1 | 1 |
| 4 | 1     | 1 | 0 | 1 | 0 | 1 |
| 5 | 1     | 1 | 0 | 1 | 0 | 0 |
| 6 | 0     | 1 | 0 | 0 | 1 | 0 |
| 7 | 1     | 1 | 1 | 1 | 1 | 1 |
| 8 | 0     | 0 | 0 | 0 | 1 | 0 |



**Este dossier está hecho para seguir la clase de prueba.**

**Si te apuntas al curso te enviaremos por correo el dossier entero con todos los temas que faltan, ejercicios y exámenes de años anteriores**

**Más información en:**

**[www.estadistix.com](http://www.estadistix.com)**

**Y si tienes cualquier consulta,  
escríbenos un whatsapp al 644310902**

