

# PRUEBAS CRITERIALES

---

Unos test que se estudiaran a parte son los criteriosales , estos test criteriosales son aquellos que se evaluan referidos a criterios, también llamados dominio de interés. Busca evaluar cuanto los sujetos dominan un criterio.

Los test tradicionales TRN se llaman test referidos a la norma porque intentan evaluar tomando como referencia a un individuo o grupo normativo, o sea buscan graduar a las personas según el grupo.

Aunque los test referidos a criterio TRC buscan evaluar si se domina un criterio , tambien sirven para clasificar a los individuos simplemente sumando las respuestas correctas.

La primera vez que se habló de ellos fue Robert Glaser en 1963 en la revista American Psychologist

Berk los define como "costruidos para evaluar el nivel de ejecución de un sujeto con respecto a un dominio de conductas bien definidas"

Características aparecieron gracias a estos TRC los definió Hambleton y dice:

a) obliga a los profesores y constructores de tests a definir con mayor claridad y operatividad los objetivos o criterios de interés, para constuir los test correspondientes para su evaluación.

b) obligan a a muestrear exhaustivamente los objetivos a evaluar , mirar la validez de contenido y que los objetivos estén bien representados.

c) Aparecieron nuevos métodos para evaluar la fiabilidad y la validez de los test, el coeficiente Kappa por ejemplo.

d)Estos test buscan casi siempre dividir en dos los grupos de origen los que dominan el criterio y los que no, para ello hay que establecer de un modo adecuado el punto de corte. Este punto lo establecerán expertos en el tema y en establecer evaluaciones. ESTE PUNTO DE CORTE SE DEBE ESTABLECER EN FUNCIÓN DE LA MATERIA Y NO DE LAS PERSONAS.

e) Estos test han ayudado mucho a diagnosticar las deficiencias de aprendizaje. Permiten detectar los puntos fuertes y débiles de los sujetos y a tomar decisiones sobre la enseñanza. (diagnóstico TDAH, dislexia...)

f) Ha obligado a los profesores a ponerse al día en procesos matemáticos de evaluación de los estudiantes. Y así se a mejorado el proceso de evaluación , siendo más justa y de calidad.

Ante todo estos TRC deben tener en cuenta dos posibles problemas con los que se encuentran :

- deben ser muy objetivos a la hora de evaluar

- se debe garantizar la validez de contenido de la prueba

## TIPOS DE PRUEBAS CRITERIALES

Las pruebas criterioles se clasifican en función de las instrucciones y pueden ser:

- test de certificación o admisión , que permiten ver si una persona supera un nivel educativo que se necesita
- test de nivel o dominio, si ya esta en el nivel y puede pasar de curso
- test de aula
- test diagnostico busca ver problemas de aprendizaje, cual es su origen y características.
- test individualizado para ver el avance de un individuo, para ver como realiza los procesos y cual es su rendimiento.

## 3 ELABORACIÓN DE PRUEBAS CRITERIALES

Nos vamos a centrar en los test criterioles de clases porque son los más usados.

### 3.1. ESPECIFICACIÓN DEL DOMINIO EDUCATIVO

Para crear una test criterial, es muy importante tener bien claro el dominio de instrucción, sobretodo para poder evaluar la validez de contenido y constructo. Las unidades dominio no deben solaparse, se debe preguntar el dominio en su totalidad. Así se debe crear ítems para todos los objetivos contenidos y actividades.

Si queremos crear ítems nos debemos ir a la clasificación de Jornet y Suárez:

Así podemos trabajar con los siguientes tipos de criterios:

Procedimientos	Formatos de ítems Transformación lingüística
Métodos basados en teorías	Diseño de facetas Conceptos, reglas y principios Aproximación de Markle y Tiemann Sistema LOGIQ Inventario de calidad Instruccional (IQI)
Procedimientos orientados al diagnóstico	Algoritmo de Scandura Métodos basados en estrategias cognitivas

El procedimiento de formatos de ítems: son aplicables cuando se puede realizar un análisis de tareas o donde se puede efectuar una definición jerarquizada del dominio de instrucción. Esto facilita luego la utilización de programas de instrucción .

El método de transformación lingüística los ítems se encuentran por medio de cuatro pasos:

- Búsqueda de frases relevantes en el proceso de instrucción

- Selección de las frases más destacadas
- Transformación de las frases según el tipo ítem, siguiendo las taxonomías de Bloom: Reconocimiento, comprensión , aplicación y análisis.
- Construcción de distractores en los formatos de elección multiple

El método de diseño de facetas: en el se transforman los objetivos operativos en sentencias directrices que constan de dos tipos de elementos:

- Formulación fija de una parte del ítem
- Formulación de una parte variables.
- La combinación de la parte fia con la parte variable da lugar a diferentes tipos de ítems.

## Conceptos reglas y principios ,

El de Tiemann y Markle crea ítems basados en el uso de ejemplos y contraejemplos de un concepto.

La metodología de inventario de calidad instruccional de IQI . que permite categorizar los objetivos e ítems , permitiendo la producción de estos últimos.

De los procedimientos orientados al diagnostico:

- El algoritmo de Scandura analiza por diagrama de flujo los procesos implicados en la resolución de un problema.
- Los métodos basados en estrategias cognitivas permiten analizar los procedimientos utilizados en la resolución de problemas o la forma de actuar de los alumnos exitosos y no exitosos.

## Analisis de ítems.

Los items se analizan desde una visión lógica y una empírica:

La revisión lógica mira si cada item trata de cada contenido y lo analizan expertos , miran lo representativo del item , su calidad, y el sesgo la neutralidad del item .

La revisión empírica mira la aplicación de la prueba y su funcionamiento , para ello habla de las etapas que son la formulación de las hipótesis , la aplicación piloto, analisis de resusltados como indice de dificultad , discriminación y validez de los items.

Interpretación de los resultados

A la hora de evaluar estos ítems tenemos que tener en cuenta que deben discriminar más que los test referidos a la norma, por lo tanto se deben hacer más pruebas de discriminación, como el índice de sensibilidad a las instrucciones que mira si hay diferencia entre el grupo de control o el experimental . Además se aplica el  $DIS_{GN}$  o sea ganancia neta. Que es la proporción de sujetos que responden incorrectamente al ítem en el pretest y correctamente en el posttest, menos la proporción de sujetos que respondieron incorrectamente al pretest y en el postes. Su valor va de -1 a 1

La  $DIS_{CI-CS}$  o de diferencia entre los grupos con instrucciones y sin instrucciones, donde restamos la proporción de los que aciertan el ítem con instrucciones los que lo aciertan sin instrucciones :

Sujetos ( caso 1) Sujetos emparejados ( caso 2)	Pretest ( caso 1) Grupo Sin Instrucción ( caso 2)		Postest ( caso 1) Grupo Con Instrucción ( caso 2)	
	Ítem		Ítem	
	1	2	1	2
1	0	1	0	1
2	0	0	1	1
3	1	0	1	1
4	0	0	1	1
5	1	0	1	1
6	0	0	0	0
7	0	1	1	1
8	1	0	1	0
9	0	0	0	1
9	0	0	0	1

**Tabla 8.3.** Índices de dificultad y discriminación del ejemplo propuesto

Sujetos ( caso 1) Sujetos emparejados ( caso 2)	Pretest ( caso 1) Grupo Sin Instrucción ( caso 2)		Postest ( caso 1) Grupo Con Instrucción ( caso 2)	
	Ítem		Ítem	
	1	2	1	2
Dificultad (pi)	0,4	0,2	0,7	0,8
$DIS_{GN}$ ( caso 1)	0,0 (3/10-3/10)	0,4		
$DIS_{CI-SI}$ ( caso 2)	0,3 (7/10-4/10)	0,6		

El índice de homogeneidad. Busca que los ítems sean parecidos , y pueden ser cualquiera de los del tema dos

### DETERMINACIÓN DE ESTANDARES Y PUNTOS DE CORTE

Es el valor que nos indica si una persona pasa o no pero en el fondo son terminos parecidos pero no iguales, el estandar es solo teórica y el punto de core es observable.

Estos estandares o puntos de corte lo eligen expertos, habiendo asignado previamente niveles de dificultad a los objetivos y niveles de relevancia .

Par jorner y Suarez se clasifican los diferentes metodos así:

<b>Tipo de objetivo métrico</b>	<b>Categorías metodológicas</b>
Métodos para la determinación de un estándar	Métodos basados en juicios sobre ítems Métodos basados en juicios sobre los sujetos Métodos basados en el juicio sobre el grupo de referencia
Métodos para el establecimiento de un punto de corte	Modelos de estado Modelos continuos basados en la teoría de la decisión Métodos de compromiso Otros métodos

### ESTUDIO DE CARACTERÍSTICAS TÉCNICAS DE LA PRUEBA .

Son la fiabilidad y la validez que mide si la prueba da resustados precisos y si mide lo que debe de medir en sucesivas pruebas. La validez debe atender a una triple dimensión el contenido , el constructo y predictiva o de criterio. Lo normal es utilizar el KAPPA COHEN.