

PSICOMETRÍA

1.	INTRODUCCIÓN A LA PSICOMETRÍA		
2.	PRINCIPIOS BÁSICOS PARA LA CONSTRUCCIÓN DE INSTRUMENTOS DE MEDICIÓN PSICOLÓGICA	}	CONSTRUCCIÓN DE INSTRUMENTOS DE MEDICIÓN PSICOLÓGICA
3.	TÉCNICAS PARA LA CONSTRUCCIÓN DE ESCALAS DE ACTITUDES		
4.	FIABILIDAD DE LAS PUNTUACIONES		
5.	FIABILIDAD DE LOS TEST REFERIDOS AL CRITERIO	}	EVALUACIÓN DE LAS PROPIEDADES MÉTRICAS DE LOS INSTRUMENTOS DE MEDICIÓN PSICOLÓGICA
6.	VALIDEZ DE LAS INFERENCIAS (I)		
7.	VALIDEZ DE LAS INFERENCIAS (II)		
8.	ANÁLISIS DE LOS ÍTEMS		
9.	ASIGNACIÓN, TRANSFORMACIÓN Y EQUIPARACIÓN DE LAS PUNTUACIONES	}	APLICACIÓN DE LOS INSTRUMENTOS Y EVALUACIÓN DE LOS SUJETOS

TEMA 1. INTRODUCCIÓN A LA PSICOMETRÍA

1. METODOLOGÍA DE LA PSICOLOGÍA COMO CIENCIA DEL COMPORTAMIENTO (ESTUDIO DE LA CONDUCTA EN SENTIDO AMPLIO)

DISEÑOS DE INVESTIGACIÓN	ANÁLISIS DE DATOS	PSICOMETRÍA
• <u>OPERATIVIZACIÓN DE LAS VARIABLES DE LA HIPÓTESIS.</u> • ELABORACIÓN DE UN PLAN DE TRABAJO O PROCEDIMIENTO PARA LA <u>RECOGIDA DE DATOS.</u>	• TÉCNICAS PARA EL <u>TRATAMIENTO ESTADÍSTICO DE LOS DATOS.</u>	• <u>REGLAS PARA EL PROCESO DE OPERATIVIZACIÓN DE LAS VARIABLES</u> QUE SE QUIEREN MEDIR. • <u>ANÁLISIS DE FIABILIDAD</u> (ERROR) Y <u>VALIDACIÓN</u> PARA PODER HACER INFERENCIAS.

* FUNDAMENTOS DE INVESTIGACIÓN ES LA 4ª ASIGNATURA QUE COMPLETA ESTE MARCO DIDÁCTICO.

2. CONCEPTO DE PSICOMETRÍA

ANÁLISIS ETIMOLÓGICO: PSYKHE + METRUM: "MEDIDA DE LA PSYKHE"

ANÁLISIS DE OTRAS DEFINICIONES:

"LA PSICOMETRÍA ES UNA DISCIPLINA METODOLÓGICA, DENTRO DEL ÁREA DE LA PSICOLOGÍA, CUYA TAREA FUNDAMENTAL ES LA MEDICIÓN O CUANTIFICACIÓN DE LAS VARIABLES PSICOLÓGICAS CON TODAS LAS IMPLICACIONES QUE ELLO CONLLEVA, TANTO TEÓRICAS (POSIBILIDADES Y CRITERIOS DE MEDICIÓN) COMO PRÁCTICAS (CÓMO Y CON QUÉ SE MIDE)".

PARA ELLO DEBERÁ:	1.	<u>JUSTIFICAR Y LEGITIMAR LA MEDICIÓN PSICOLÓGICA:</u>
	1.	DESARROLLAR MODELOS FORMALES (PARA REPRESENTAR LOS FENÓMENOS A ESTUDIAR: HECHOS → DATOS)
	2.	VALIDAR LOS MODELOS DESARROLLADOS (EN QUÉ MEDIDA REPRESENTAN LA REALIDAD)
	3.	ESTABLECER LAS CONDICIONES PARA EL PROCESO DE MEDIDA.
	2.	<u>OCUPARSE DE LAS IMPLICACIONES PRÁCTICAS Y APLICADAS QUE DICHA MEDICIÓN CONLLEVA:</u>
	4.	CONSTRUIR Y EVALUAR LOS INSTRUMENTOS DE MEDICIÓN.

• EL OBJETIVO FINAL DE LA PSICOLOGÍA ES LA DESCRIPCIÓN, EXPLICACIÓN Y PREDICCIÓN DE LOS FENÓMENOS PSICOLÓGICOS. COMO CIENCIA AVANZADA QUE ES, TENDRÁ SU BASE EN LA MEDICIÓN, QUE LE PERMITIRÁ CONTRASTAR EMPÍRICAMENTE LAS HIPÓTESIS PLANTEADAS PARA ESTABLECER SUS LEYES GENERALES.

• LA PSICOMETRÍA CONTRIBUYE, POR TANTO, A FUNDAMENTAR, ELABORAR Y CONTRASTAR TODAS LAS TEORÍAS PSICOLÓGICAS.

3. MEDICIÓN

PARA NUNNALLY (1970)	<u>CONJUNTO DE NORMAS PARA ASIGNAR NÚMEROS A LOS OBJETOS</u> . ESTOS NÚMEROS REPRESENTAN <u>CANTIDADES DE ATRIBUTOS</u> (CARACTERÍSTICAS DE LOS OBJETOS (ESTÍMULOS, SUJETOS Y RESPUESTAS) Y NO LOS OBJETOS MISMOS)
PARA ZELLER Y CARMINES (1980)	PROCESO MEDIANTE EL CUAL SE ENLAZAN <u>CONCEPTOS ABSTRACTOS</u> (<u>CONSTRUCTOS</u> INOBSERVABLES DIRECTAMENTE, TEÓRICOS, VARIABLES LATENTES), CON <u>INDICADORES EMPÍRICOS OBSERVABLES DIRECTAMENTE</u> (<u>CONDUCTAS</u>) * <u>MEDICIÓN POR INDICADORES</u> : SELECCIONAR UNA SERIE DE <u>INDICADORES QUE SÍ PUEDEN SER MEDIDOS DIRECTAMENTE</u> Y QUE SE SUPONE QUE ESTÁN EN ESTRECHA RELACIÓN CON EL CONSTRUCTO O VARIABLE PSICOLÓGICA QUE SE QUIERE MEDIR.
OBJETIVOS FUNDAMENTALES: ESTIMAR LOS <u>ERRORES ALEATORIOS</u> (<u>FIABILIDAD DE LAS MEDIDAS</u>) Y GARANTIZAR QUE SIRVE PARA EXPLICAR Y PREDECIR LOS <u>FENÓMENOS DE INTERÉS</u> (<u>VALIDEZ DE LAS MEDIDAS</u>).	

UNIDADES DE MEDIDA

<u>REFERIDA A NORMAS</u> LA MÁS HABITUAL. CONSISTE EN COMPARAR LOS RESULTADOS OBTENIDOS POR EL SUJETO CON LOS OBTENIDOS POR UN GRUPO DE REFERENCIA (<u>GRUPO NORMATIVO</u>) (MISMA POBLACIÓN, MISMA CLASE, EDAD, VARONES, MUJERES, NIVEL DE ESTUDIOS,...)	<u>REFERIDA AL CRITERIO</u> SE HACE EN FUNCIÓN DE LA SUPERACIÓN O NO DE UN <u>PUNTO CRÍTICO</u> QUE VA A DAR SIGNIFICADO A LAS PUNTUACIONES OBTENIDAS (EXÁMENES)
--	---

4. ORÍGENES Y DESARROLLO

HAY VARIOS <u>HECHOS FUNDAMENTALES</u> QUE CONTRIBUYEN A LA CONVERSIÓN DE LA PSICOLOGÍA EN UNA CIENCIA POSITIVA INDEPENDIENTE DE LA FILOSOFÍA: (1) LA TENDENCIA A FORMULAR LOS PROBLEMAS CIENTÍFICOS EN TÉRMINOS MATEMÁTICOS Y, (2) EL ENFRENTAMIENTO, EN EL SIGLO XIX, CON DOS PROBLEMAS FUNDAMENTALES : EL PSICOFÍSICO Y LA CUANTIFICACIÓN DE LAS DIFERENCIAS INDIVIDUALES.	
LA PSICOMETRÍA SURTIÓ HACIA LA <u>MITAD DEL SIGLO XIX</u> Y SE DESARROLLÓ A TRAVÉS DE: <u>LOS ESTUDIOS DE PSICOFÍSICA</u> : - DIERON LUGAR A LOS <u>MÉTODOS DE ESCALAMIENTO DE ESTÍMULOS</u> (asignar valores numéricos a éstos: escalas psicofísicas y psicológicas. Fechner, Stevens y Thurstone) - DESARROLLO DE LA <u>INVESTIGACIÓN EXPERIMENTALISTA</u>	<u>LOS ESTUDIOS DE LAS DIFERENCIAS INDIVIDUALES</u> : - DIERON LUGAR AL <u>MÉTODO DE LOS TESTS PARA EL ESCALAMIENTO DE SUJETOS</u> (Calton, Pearson, Cattell, Binet, Terman: pioneros en la medida de las diferencias individuales) - DESARROLLO DE LA <u>INVESTIGACIÓN CORRELACIONAL</u>
* TAMBIÉN EXISTE EL <u>ESCALAMIENTO DE RESPUESTAS</u> , QUE ES EL <u>ESCALAMIENTO SIMULTÁNEO DE ESTÍMULOS Y SUJETOS</u> .	

5. MÉTODOS DE ESCALAMIENTO

SU <u>OBJETIVO FUNDAMENTAL</u> ES LA <u>CONSTRUCCIÓN DE ESCALAS DE MEDIDA</u> .			
DOS SUPUESTOS BÁSICOS:	1. LA EXISTENCIA DE UN <u>CONJUNTO LATENTE O SUBYACENTE</u> , A LO LARGO DEL CUAL VARÍAN LOS OBJETOS PSICOLÓGICOS A ESCALAR (ESTÍMULOS, SUJETOS, RESPUESTAS), QUE NO PUEDE SER OBSERVADO DE FORMA DIRECTA. 2. QUE LOS OBJETOS PSICOLÓGICOS <u>PUEDEN SITUARSE DE FORMA ORDENADA</u> A LO LARGO DE ESE CONTINUO.		
ESCALAMIENTO DE ESTÍMULOS	ESCALAMIENTO DE RESPUESTAS	ESCALAMIENTO DE SUJETOS	
- EL OBJETIVO ES <u>ESCALAR LOS ESTÍMULOS AMBIENTALES</u> QUE INCIDEN SOBRE EL ORGANISMO PRODUCIENDO UNA RESPUESTA. OBTENEMOS UNA ESCALA ORDENADA EN FUNCIÓN DE LOS VALORES ESCALARES DE LOS ESTÍMULOS (VALOR NUMÉRICO ASIGNADO) - LOS <u>SUJETOS</u> ACTÚAN COMO <u>INSTRUMENTOS DE MEDIDA</u> YA QUE EMITEN JUICIOS ACERCA DE LAS CARACTERÍSTICAS PERCIBIDAS EN LOS ESTÍMULOS. <u>RESPUESTA SUBJETIVA</u> QUE NOS PERMITE DIFERENCIAR Y ESCALAR.	- EL OBJETIVO ES <u>ESCALAR TANTO LOS ESTÍMULOS COMO LOS SUJETOS</u> SOBRE UN MISMO CONTINUO ORIGINANDO UNA <u>ESCALA DE ENTRELAZAMIENTO</u>. <u>ESTÍMULOS Y SUJETOS</u> SON <u>INSTRUMENTOS DE MEDIDA</u>. <u>APROXIMACIÓN CENTRADA EN LA RESPUESTA</u>. LAS VARIACIONES EN LA R SE ATRIBUYEN TANTO A LOS ESTÍMULOS COMO A LOS SUJETOS (<u>LA POSICIÓN DE SUJETO RESPECTO DEL ATRIBUTO QUE SE ESTÁ ANALIZANDO ES UN FACTOR DETERMINANTE</u>).	- EL OBJETIVO ES <u>ESCALAR A LOS SUJETOS</u> A LO LARGO DE UN CONTINUO PARA ESTUDIAR LAS <u>DIFERENCIAS INDIVIDUALES</u>. PROPORCIONA UNA PUNTUACIÓN NUMÉRICA PARA CADA SUJETO DE LA MUESTRA. - LOS <u>ESTÍMULOS</u> (ÍTEMS) ACTÚAN COMO <u>INSTRUMENTOS DE MEDIDA</u>. - LAS DIFERENCIAS ENCONTRADAS EN LAS PUNTUACIONES OBTENIDAS REFLEJAN LAS <u>DIFERENCIAS ENTRE LOS SUJETOS RESPECTO AL ATRIBUTO O CARACTERÍSTICA QUE SE ESTÁ MIDIENDO</u>.	
ESCALAMIENTO PSICOFÍSICO (WEBER Y FECHNER)	ESCALOGRAMA DE GUTTMAN	TEORÍA CLÁSICA DEL TEST _ TCT TEORÍA DE RESPUESTA AL ÍTEM _ TRI TEORÍA DE LA GENERALIZABILIDAD _ TG	
ESCALAMIENTO PSICOLÓGICO (THURSTONE)			

ESCALAMIENTO PSICOLÓGICO (LO VEMOS EN EL TEMA 3)

THURSTONE	ESCALAMIENTO DE ESTÍMULOS CON MODELOS MATEMÁTICOS	LEY DEL JUICIO COMPARATIVO	MÉTODO DE LAS COMPARACIONES BINARIAS.
		LEY DEL JUICIO CATEGÓRICO	MÉTODO DE LOS INTERVALOS SUCESIVOS Y DE LOS INTERVALOS APARENTEMENTE IGUALES Y MÉTODO DE ORDENACIÓN DE RANGOS.
GUTTMAN	ESCALAMIENTO DE RESPUESTAS	MÉTODO DEL ESCALOGRAMA	ESCALA DE ENTRELAZAMIENTO
COOMBS		TEORÍA DEL DESPLIEGUE	

7. ORÍGENES Y DESARROLLO DE LOS TEST (ESCALAMIENTO DE SUJETOS)

TEST: ES UN INSTRUMENTO DE MEDICIÓN DISEÑADO PARA ESTUDIAR DE UN MODO OBJETIVO Y SISTEMÁTICO EL NIVEL DE LOS SUJETOS RESPECTO A ALGÚN ATRIBUTO, CARACTERÍSTICA O DOMINIO DE CONOCIMIENTOS Y, A PARTIR DE LAS PUNTUACIONES QUE OBTENGAN LOS SUJETOS EN EL TEST, PODEMOS ANALIZAR LAS DIFERENCIAS INDIVIDUALES.

TRES FACTORES DECISIVOS EN EL DESARROLLO DE LOS TEST:	1. <u>APERTURA DEL LABORATORIO ANTROPOMÉTRICO DE GALTON:</u> ESTUDIO DE LA RELACIÓN ENTRE LAS CARACTERÍSTICAS MENTALES Y LAS FÍSICAS.
	2. <u>DESARROLLO DE LA CORRELACIÓN DE PEARSON</u>
	3. <u>INTERPRETACIÓN DE SPEARMAN</u> DE DICHA CORRELACIÓN ENTRE VARIABLES CONSIDERANDO QUE DETERMINA LA EXISTENCIA DE UN FACTOR COMÚN.

LA IMPORTANCIA DE LOS TRABAJOS DE GALTON

GALTON	PRIMEROS TEST PSICOMÉTRICOS	- COMPRENDIÓ QUE PARA OBTENER UNA DESCRIPCIÓN COMPLETA DEL HOMBRE HABÍA QUE COMPLETAR LAS MEDIDAS ANTROPOMÉTRICAS CON LAS PSICOLÓGICAS. RECOGIÓ NUMEROSOS DATOS Y COMENZÓ A BUSCAR LAS RELACIONES ENTRE ELLOS. FUE EL PRIMERO EN APLICAR CONCEPTOS ESTADÍSTICOS DE DISTRIBUCIÓN NORMAL, DE MEDIA, DE MEDIANA, VARIANZA Y CORRELACIÓN A DATOS PSICOLÓGICOS.
		- PARA LA MEDICIÓN CONSTRUYÓ Y UTILIZÓ UNA SERIE DE INSTRUMENTOS QUE PUEDEN SER CONSIDERADOS COMO TALES (CARÁCTER ANTROPOMÉTRICO, SENSORIAL Y MOTOR).

* MIENTRAS WUNDT TRATABA DE ENCONTRAR LAS LEYES GENERALES QUE DIRIGEN Y GOBIERNAN LA MENTE HUMANA, GALTON (INFLUENCIAS DE LOS EVOLUCIONISTAS) TRATABA DE CLASIFICAR Y CATALOGAR LAS DIFERENTES FORMAS DE MENTES QUE PODÍAN EXISTIR. SEPARACIÓN ENTRE LA PSICOLOGÍA EXPERIMENTAL Y DIFERENCIAL.

LOS PRIMEROS TEST MENTALES

CATTELL 1890	TEST MENTAL	- FUE EL PRIMERO EN UTILIZAR EL TÉRMINO. PARA ÉL LOS TESTE ERAN UN SISTEMA UNIFORME QUE PERMITE COMPARAR Y COMBINAR, EN LUGARES Y MOMENTOS DIFERENTES, LA MEDIDA DE LAS FUNCIONES MENTALES, PERO SUS TEST TENÍAN CARÁCTER SENSORIAL Y MOTOR. CREÍA QUE EL FUNCIONAMIENTO INTELECTUAL SE PODÍA MEDIR CON TEST DE DISCRIMINACIÓN SENSORIAL Y MIDIENDO EL TIEMPO DE REACCIÓN.
		- MÁS TARDE EL ANÁLISIS DE LOS DATOS PUSO DE MANIFIESTO QUE LA CORRELACIÓN ENTRE ESTE TIPO DE PRUEBAS Y EL NIVEL INTELECTUAL ERA NULA.

DESARROLLO DE LOS PRIMEROS TEST DE INTELIGENCIA Y POSTERIORES (WESCHLER)

ALFRED BINET 1905	ESCALA DE BINET-SIMON	- SE CENTRÓ EN LOS <u>PROCESOS MENTALES SUPERIORES</u> E INTRODUJO LA VARIABLE <u>EDAD</u>. - CREÓ LA PRIMERA ESCALA DE INTELIGENCIA JUNTO CON SIMON PARA MEDIR LA CAPACIDAD DE COMPRENSIÓN Y RAZONAMIENTO DE LOS NIÑOS CON 30 ELEMENTOS DE DIFICULTAD CRECIENTE (LA MAYORÍA VERBALES) - SE DESARROLLA EL CONCEPTO DE <u>EDAD MENTAL</u>: <u>GRUPOS NORMATIVOS</u> (PUNTUACIÓN MEDIA DE NIÑOS INTELECTUALMENTE NORMALES POR EDAD)
TERMAN 1916	TEST DE STANFORD-BINET	- <u>ADAPTACIÓN DEL TEST DE BINET</u>. SIGUE SU TÁCTICA PARA LA ELABORACIÓN DE TEST. - INTELIGENCIA: HABILIDAD PARA LLEVAR A CABO RAZONAMIENTOS ABSTRACTOS. - UTILIZA EL CONCEPTO DE <u>COCIENTE INTELECTUAL</u> (STERN, 1912) COMO MEDIDA DE LA INTELIGENCIA (PRIMERA ESCALA DE PUNTUACIONES REFERIDA A APTITUDES MENTALES): $CI = \frac{EM}{EC}$
WESCHLER 1938 (1955) 1949 1967	TEST DE WESCHLER-BELLEVUE (WAIS) TEST WISC WESCHLER PRESCHOOL AND PRIMARY SCALE OF INTELLIGENCE	- MEDICIÓN DE LA INTELIGENCIA PARA <u>ADULTOS</u> (WESCHLER ADULT INTELLIGENCE SCALE) - VERSIONES PARA <u>NIÑOS</u>

LOS TEST COLECTIVOS

R. M. YERKES ARTHUR OTIS 1918	TEST ALPHA Y BETA	LA NECESIDAD DE APLICAR TEST EN EL <u>EJÉRCITO</u> (1ªGM) LLEVÓ A SU CREACIÓN. - TEST ALPHA: PARA LA POBLACIÓN GENERAL. - TEST BETA: PARA RECLUTAS ANALFABETOS O SIN DOMINIO DEL INGLÉS.
EN 1922 CATELL FUNDÓ LA <u>PRIMERA EMPRESA DEDICADA A LA PUBLICACIÓN MASIVA DE TEST</u> , A PARTIR DE AHÍ EMPEZARON A USARSE COMO <u>INSTRUMENTO DE SELECCIÓN EN LA ADMINISTRACIÓN AMERICANA</u> .		

LAS BATERÍAS DE APTITUD MÚLTIPLE

SE DESARROLLA EL COEFICIENTE DE CORRELACIÓN DE <u>PEARSON</u> Y EL ANÁLISIS FACTORIAL DE <u>SPEARMAN</u> (TÉCNICAS CORRELACIONALES) QUE NOS APORTAN: - <u>FUNDAMENTACIÓN TEÓRICA DEL FACTOR GENERAL G</u>. - NECESIDAD DE ESTUDIAR OTRAS <u>APTITUDES MÁS ESPECÍFICAS</u> EN EL CAMPO DE LA INTELIGENCIA. SURGEN LAS BATERÍAS DE APTITUD MÚLTIPLE PARA OBTENER UNA PUNTUACIÓN POR SEPARADO DE CADA RASGO, EN LUGAR DE UNA PUNTUACIÓN TOTAL (CI)		
THURSTONE 1938 (1941)	BATERÍA DE APTITUDES MENTALES PRIMARIAS	MEDIDA DE <u>COMPONENTES FUNDAMENTALES</u> DE LA INTELIGENCIA: <u>COMPRENSIÓN VERBAL, FLUIDEZ VERBAL, APTITUD NUMÉRICA, APTITUD ESPACIAL, MEMORIA, RAPIDEZ PERCEPTIVA Y RAZONAMIENTO GENERAL</u> .
	DAT, PMA	BATERÍAS ACTUALES DE USO HABITUAL.

LOS TEST DE PERSONALIDAD

KRAEPELIN (1894)	TEST DE ASOCIACIÓN LIBRE	- PRECURSOR DE LOS TEST DE PERSONALIDAD
WOODWORTH (1917)	HOJA DE DATOS PERSONALES	- PROTOTIPO DE CUESTIONARIO DE PERSONALIDAD CON CARÁCTER PSICOMÉTRICO, UTILIZADO EN LA 1ªGM. - DETECCIÓN DE SUJETOS CON <u>TRASTORNOS NEURÓTICOS GRAVES</u>.

HATHAWAY Y MCKINLEY (1942)	MMPI INVENTARIO MULTIFASICO DE PERSONALIDAD DE MINNESOTA	- PARA POBLACIONES CLÍNICAS . - ORIENTADO HACIA LA CLASIFICACIÓN PSIQUIÁTRICA.
GUILFORD (1949)	TEST DE PERSONALIDAD	- PARA POBLACIONES NORMALES . - ORIENTADO HACIA LOS RASGOS.
CATTELL (1970)		
RORSCHACH 1921	T DE PSICODIAGNÓSTICO O TEST DE LAS MANCHAS DE TINTA	- SON TEST PROYECTIVOS . ASOCIADOS A LA PSICOLOGÍA CLÍNICA, ESPECIALMENTE AL PSICOANÁLISIS. - HAY VARIAS RAZONES QUE HAN CONTRIBUIDO A LA DISMINUCIÓN DEL USO DE ESTAS TÉCNICAS . • AUMENTO DE LA SOFISTICACIÓN CIENTÍFICA • PROCEDIMIENTOS ALTERNATIVAS: MMPI • ESTÁN ASOCIADAS A LA TEORÍA PSICOANALÍTICA.
MACHOVER 1926	TEST DE DIBUJO DE LA FIGURA HUMANA	
1938	TEST DE APERCEPCIÓN TEMÁTICA TAT	- A PESAR DE TODO ESTO EL TEST DE RORSCHACH SIGUE MOSTRANDO SU UTILIDAD.

LA MEDICIÓN DE LOS INTERESES Y ACTITUDES

STRONG (1927)	CUESTIONARIO DE INTERESES VOCACIONALES (SVIB)	LIKERT (1932)	TÉCNICA DE LIKERT
KUDER (1934/85)	ESCALA DE PREFERENCIAS DE KUDER	OSGOOD (1952)	EL DIFERENCIAL SEMÁNTICO
THURSTONE (1928)	LEY DEL JUICIO COMPARATIVO Y LEY DEL JUICIO CATEGÓRICO	GUTTMAN	EL ESCALOGRAMA

INSTITUCIONALIZACIÓN DEL USO DE LOS TEST

EL USO MASIVO DE TEST FUE EXTENDIENDO SU APLICACIÓN Y SE HIZO NECESARIA LA EXISTENCIA DE INSTITUCIONES TANTO PÚBLICAS COMO PRIVADAS DEDICADAS A SU GESTIÓN:	
1928	CEEB COLLEGE ENTRANCE EXAMINATION BOARD . SELECCIÓN DE ESTUDIANTES PARA ACCESO A SECUNDARIA. • EN PRINCIPIO: TEST DE CONTENIDOS CURRICULARES. • EN 1928-29: SAT SCHOLASTIC APTITUDE TEST: MEDICIÓN DE LA CAPACIDAD DE APRENDIZAJE.
1947	ETS (EDUCATIONAL TESTING SERVICE) : SE UNIERON VARIAS INSTITUCIONES (ENTRE ELLAS CEEB Y AMERICAN COUNCIL ON EDUCATION) PARA POTENCIAR LA INVESTIGACIÓN EN EL CAMPO EDUCATIVO.
1951	APA (AMERICAN PSYCHOLOGICAL ASSOCIATION) ; AERA (AMERICAN EDUCATIONAL RESEARCH ASSOCIATION) ; NCME (NATIONAL COUNCIL ON MEASUREMENTS USED IN EDUCATION) : RECOMENDACIONES TÉCNICAS PARA EL USO DE LOS TEST. EN 1999: STANDARDS FOR EDUCATIONAL AND PSYCHOLOGICAL TESTING . ÚLTIMA EDICIÓN EN 2014. EN 1999, APA, AERA Y NCME PUBLICARON: STANDARDS FOR EDUCATIONAL AND PSYCHOLOGICAL TESTING (CRITERIOS PARA EVALUACIÓN DE LOS TEST, ÚLTIMA EDICIÓN DE 2014)
1995 (ESPAÑA)	COMISIÓN DE TEST POR EL COLEGIO OFICIAL DE PSICÓLOGOS . QUE FUNCIONA COORDINADA CON OTRAS COMISIONES INTERNACIONALES COMO TASK FORCE (FEDERACIÓN EUROPEA DE ASOCIACIONES PROFESIONALES DE PSICÓLOGOS, EFPPA) O LA INTERNATIONAL TESTS COMMISSION (ITC) .

LOS TEST REFERIDOS AL CRITERIO FRENTE A LOS REFERIDOS A NORMAS

SURGEN A FINALES DE LOS 60 ANTE LA NECESIDAD DE EVALUAR EL RENDIMIENTO Y HACER UN DIAGNÓSTICO DE LAS NECESIDADES DE LOS ALUMNOS , DE CARA A UNA POSIBLE INTERVENCIÓN.	
TEST REFERIDOS A LAS NORMAS TRN	TEST REFERIDOS AL CRITERIO TRC
DESARROLLADOS PARA EVALUAR EL GRADO DE APTITUD O EL NIVEL DE LOS SUJETOS EN UN DETERMINADO RASGO EN RELACIÓN AL GRUPO NORMATIVO .	SE UTILIZA PARA EVALUAR EL ESTATUS ABSOLUTO DEL SUJETO CON RESPECTO A ALGÚN DOMINIO DE CONDUCTAS BIEN DEFINIDO (POPHAM)
DEPENDEN DEL ESTATUS DEL GRUPO (GLASER)	DEPENDEN DEL ESTATUS ABSOLUTO DE CALIDAD DEL SUJETO (GLASER)
SU OBJETIVO ES PONER DE MANIFIESTO LAS DIFERENCIAS INDIVIDUALES EN EL RASGO O CONDUCTA QUE MIDE EL TEST.	EL OBJETIVO ES ESTIMAR EL RENDIMIENTO O CONDUCTA DEL SUJETO EN LOS OBJETIVOS QUE MIDE EL TEST.
NO SE TIENE UNA DELIMITACIÓN CLARA DEL DOMINIO DE CONTENIDOS QUE SE VAN A EVALUAR	SE ESPECIFICA CLARAMENTE EL DOMINIO DE CONTENIDOS O CONDUCTAS A EVALUAR Y EL USO QUE SE VA A HACER DEL TEST.
LOS ÍTEMS DEBEN PONER EN RELIEVE LAS DIFERENCIAS INDIVIDUALES MAXIMIZANDO LA VARIANZA DEL TEST , SELECCIONANDO ÍTEMS DE DIFICULTAD MEDIA Y ALTO PODER DISCRIMINANTE .	LOS ÍTEMS SE SELECCIONAN EN FUNCIÓN DE LOS OBJETIVOS Y DEL USO DEL TEST .
LA PUNTUACIÓN OBTENIDA SE CONSIDERA UN INDICADOR DE SU PUNTUACIÓN VERDADERA EN UN RASGO LATENTE.	LA PUNTUACIÓN OBTENIDA ES UN ESTIMADOR DE LA CONDUCTA O RENDIMIENTO DEL SUJETO EN EL DOMINIO
LA PUNTUACIÓN TIENE SIGNIFICADO ÚNICAMENTE CON RELACIÓN A LOS RESULTADOS DEL GRUPO NORMATIVO .	LAS PUNTUACIONES TIENEN SIGNIFICADO EN TÉRMINOS ABSOLUTOS .
LA INTERPRETACIÓN DE UN SUJETO SE HACE COMPARÁNDOLA CON LA OBTENIDA POR EL GRUPO.	LA INTERPRETACIÓN SE HACE EN RELACIÓN A UN DOMINIO BIEN DEFINIDO MÁS QUE EN RELACIÓN A UN GRUPO NORMATIVO.

LOS TEST ADAPTATIVOS INFORMATIZADOS (TAI's)

- LA SELECCIÓN DE LOS ÍTEMS SE VA HACIENDO A LO LARGO DEL PROCESO DE ADMINISTRACIÓN DEL TEST (ADAPTÁNDOSE AL NIVEL DEL SUJETO)
- SE UTILIZA EL ORDENADOR COMO MEDIO DE SELECCIÓN DE LOS ÍTEMS.

8. DESARROLLO DE LA TEORÍA DE LOS TEST

COMO CONSECUENCIA DEL DESARROLLO DEL MÉTODO DE LOS TEST, SE HACE NECESARIO UN MARCO TEÓRICO QUE POSIBILITE LA VALIDACIÓN DE LAS INFERENCIAS Y LA ESTIMACIÓN DE LOS ERRORES DE MEDIDA (VALIDEZ Y FIABILIDAD).

LA TEORÍA DE LOS TEST PERMITE ESTABLECER UNA **RELACIÓN FUNCIONAL** (MATEMÁTICA) **ENTRE LAS VARIABLES OBSERVABLES** (PUNTUACIONES EMPÍRICAS) Y **LAS VARIABLES INOBSERVABLES** (PUNTUACIONES VERDADERAS). ESTA RELACIÓN DEBERÁ RELACIONAR EL RASGO, CONSTRUCTO O CARACTERÍSTICA CON LAS PUNTUACIONES EMPÍRICAS OBTENIDAS.

LA TEORÍA CLÁSICA DE LOS TEST (TCT)	• SE DESARROLLÓ GRACIAS A LAS APORTACIONES DE GALTON, PEARSON Y SPEARMAN. • TRES CONCEPTOS BÁSICOS: PUNTUACIÓN EMPÍRICA (X), PUNTUACIÓN VERDADERA (V) Y PUNTUACIÓN ERROR (E). • MODELO DE SPEARMAN (LINEAL Y ADITIVO): $X = V + E$ (SUS SUPUESTOS SE DESARROLLAN EN EL TEMA 4) • E INCLUYE <u>TODOS LOS ERRORES ALEATORIOS</u> QUE AFECTAN A LAS PUNTUACIONES EMPÍRICAS. ASÍ PODEMOS HACER UNA ESTIMACIÓN.
TEORÍA DE LA GENERALIZABILIDAD (TG) – <u>CRONBACH</u>	• SE CONSIDERA UNA EXTENSIÓN DE LA TCT. • ES EL MÁS AMBICIOSO Y GLOBAL INTENTO DE <u>ANÁLISIS DE TODAS LAS POSIBLES FUENTES DE ERROR</u>: <u>FACTORES INDIVIDUALES, SITUACIONALES, CARACTERÍSTICAS DEL EVALUADOR Y VARIABLES INSTRUMENTALES</u>. INTENTA DIFERENCIARLAS MEDIANTE LA APLICACIÓN DE LOS PROCEDIMIENTOS CLÁSICOS DE ANÁLISIS DE VARIANZA (AVAR) • LAS FUENTES DE VARIACIÓN (FACTORES EN TÉRMINOS DE ANOVA) SE DENOMINAN FACETAS Y LOS NIVELES DE CADA FACTOR CONDICIONES. • EL SURGIMIENTO DE LA TRI LA RELEGÓ A UN SEGUNDO PLANO.
TEORÍA DE RESPUESTA AL ÍTEM (TRI) - <u>LORD</u>	• <u>CURVA CARACTERÍSTICA DEL ÍTEM</u>: MODELOS QUE ASUMEN UNA RELACIÓN FUNCIONAL ENTRE LOS VALORES DE LA VARIABLE QUE MIDEN LOS ÍTEMS (NIVEL DE HABILIDAD DE LOS SUJETOS) Y LA PROBABILIDAD DE QUE LOS SUJETOS, EN FUNCIÓN DE SU NIVEL DE HABILIDAD, ACIERTEN CADA ÍTEM.

TEMA 2. PRINCIPIOS BÁSICOS PARA LA CONSTRUCCIÓN DE INSTRUMENTOS DE MEDICIÓN PSICOLÓGICA

CONSTRUCCIÓN DE TEST COMO INSTRUMENTO CIENTÍFICO DE MEDICIÓN. ETAPAS

INICIAL CONSTRUCCIÓN TEMA 2 Y 3	- ELABORACIÓN DE LA <u>PRUEBA PILOTO</u> - APLICACIÓN A UNA MUESTRA DE SUJETOS - ASIGNACIÓN DE PUNTUACIONES A LOS MISMOS	DETERMINAR LA FINALIDAD DEL TEST, SUS CARACTERÍSTICAS, LONGITUD, CARACTERÍSTICAS PSICOMÉTRICAS DE LOS ÍTEMS, REDACCIÓN DE LOS MISMOS, REVISIÓN CRÍTICA, CONFECCIÓN DE LA PRUEBA.
SEGUNDA EVALUACIÓN	- EVALUACIÓN DE LA <u>CALIDAD PSICOMÉTRICA</u> DE CADA UNO DE LOS <u>ÍTEMS</u> (TEMA 8) - ANÁLISIS SUBJETIVO POR UN GRUPO DE EXPERTOS O JUECES - ANÁLISIS OBJETIVO BASADO EN LAS RESPUESTAS - EVALUACIÓN DEL <u>CONJUNTO DEL TEST</u> UNA VEZ ELIMINADOS LOS ÍTEMS INADECUADOS (FIABILIDAD Y VALIDEZ) (TEMAS 4, 5, 6 Y 7) CONSTRUCCIÓN DEFINITIVA DEL TEST	
TERCERA (TEMA 9) APLICACIÓN Y EVALUACIÓN SUJETOS	- APLICACIÓN DEL TEST A UNA MUESTRA REPRESENTATIVA DE LA POBLACIÓN A LA QUE VA DIRIGIDO - ASIGNACIÓN DE PUNTUACIONES PARA LA EVALUACIÓN DE LOS SUJETOS - ESTANDARIZACIÓN DE PUNTUACIONES Y ESTABLECIMIENTO DE NORMAS PARA SU INTERPRETACIÓN	
FINAL (TEMA 9) MANUAL DEL TEST	ELABORACIÓN DEL MANUAL DEL TEST	

1. LOS TEST, ESCALAS, CUESTIONARIOS E INVENTARIOS

TEST	VARIABLES <u>COGNITIVAS</u> APTITUDES, CONOCIMIENTO, HABILIDADES, RENDIMIENTO, ETC. LAS <u>RESPUESTAS SON CORRECTAS O INCORRECTAS.</u>	<u>PUNTUACIÓN TOTAL: SUMA DIRECTA O PONDERADA</u> DE TODAS LAS RESPUESTAS CORRECTAS. DA LUGAR A UNA <u>ESCALA JERÁRQUICA O ACUMULATIVA.</u>	¿Cuál de las siguientes respuestas es mayor? (Ejemplo EFAI – 4) a) 2² b) 3 x 2 c) 5 – 3 d) 8: 4 La capital de España es: a) París b) Londres c) Madrid
ESCALAS DE CATEGORÍAS O DE CLASIFICACIÓN.		<u>PUNTUACIÓN TOTAL: SUMA DE LAS PUNTUACIONES ASIGNADAS A LAS CATEGORÍAS</u> ELEGIDAS POR EL SUJETO. LA ELECCIÓN SE HACE SOBRE UNA <u>ESCALA DE CATEGORÍAS GRADUADA Y ORDENADA.</u> MOSTRAMOS NUESTRA <u>POSICIÓN RESPECTO</u> A ALGO.	Barómetro correspondiente al mes de Noviembre de 2015 (CIS) P.1 Para empezar, refiriéndonos a la <u>situación económica</u> general de España, ¿cómo la calificaría Ud.: muy buena, buena, regular, mala o muy mala? - Muy buena 1 - Buena 2 - Regular 3 - Mala 4 - Muy mala 5 (32)
CUESTIONARIOS INSTRUMENTO USADO GENERALMENTE PARA LLEVAR A CABO UNA INVESTIGACIÓN MEDIANTE ENCUESTAS.	VARIABLES <u>NO COGNITIVAS.</u> <u>CONTENIDO AFECTIVO U ORÉCTICO.</u> ACTITUDES, INTERESES, PREFERENCIAS, OPINIONES, ETC. LAS <u>RESPUESTAS NO SON CORRECTAS NI INCORRECTAS.</u>	<u>PUNTUACIÓN TOTAL: LOS ÍTEMS SE PUEDEN PUNTUAR E INTERPRETAR INDIVIDUALMENTE.</u> LAS OPCIONES DE <u>RESPUESTAS NO ESTÁN ORDENADAS NI GRADUADAS.</u> LOS ÍTEMS SON INDEPENDIENTES. SUELEN SER MUY VARIADOS. ELEGIMOS LA/S ALTERNATIVA/S QUE PREFIRAMOS. MOSTRAMOS <u>PREFERENCIA U OPINIÓN.</u>	En esta prueba se pide que indiques tus gustos y preferencias respecto a las actividades y profesiones que se presentan en las páginas siguientes. Señala A para “Me gusta”; B para “Me es indiferente o tengo dudas”; C para “No me gusta”; y D para “No conoces esa actividad o profesión” (Ejemplo: IPP – R Intereses y Preferencias Profesionales) Apagar incendios o prevenirlos, proteger o salvar a las personas en los mismos A B C D
INVENTARIOS VINCULADOS A INSTRUMENTOS ELABORADOS PARA MEDIR VARIABLES DE PERSONALIDAD.		MOSTRAMOS CONFORMIDAD O NO CON LOS ENUNCIADOS. <u>GRADO DE ACUERDO O DESACUERDO.</u>	Este test contiene una serie de frases. Lea cada una de ellas y decida si, aplicada a Vd mismo, es verdadera o falsa (Ejemplo: MMPI-2-RF Inventario Multifásico de Personalidad de Minnesota) E1 Me gusta la música V F E2 Me levanto muy temprano V F

2. CONSTRUCCIÓN DE UN TEST: ETAPAS

DETERMINAR LA FINALIDAD DEL TEST	¿QUÉ SE VA A MEDIR?	VARIABLE OBJETO DE ESTUDIO LO PRIMERO ES DEFINIR LOS CONSTRUCTOS (VARIABLES TEÓRICAS INOBSERVABLES) PARA DETERMINAR LAS CONDUCTAS REPRESENTATIVAS DE LOS MISMOS (YA OBSERVABLES)		
	¿A QUIÉN SE VA A MEDIR?	POBLACIÓN A QUIEN VA DIRIGIDO - PARA UN MISMO CONSTRUCTO EXISTEN CONDUCTAS COMUNES A TODAS LAS POBLACIONES Y CONDUCTAS ESPECÍFICAS DE CADA POBLACIÓN. ADECUARSE. - TAMBIÉN CONTENIDO DE LOS ÍTEMS, REDACCIÓN, LONGITUD DEL TEST E INSTRUCCIONES DE APLICACIÓN Y CUMPLIMENTACIÓN SERÁN DIFERENTES.		
	¿PARA QUÉ SE VA A MEDIR?	UTILIZACIÓN PREVISTA • SELECCIÓN (ÁMBITO EDUCATIVO/PROFESIONAL) • CLASIFICACIÓN/COLOCACIÓN (ASIGNAR A GRUPOS POR NIVEL) • DIAGNÓSTICO • CERTIFICACIÓN (ACREDITAR CUALIFICACIÓN O COMPETENCIA) • ORIENTACIÓN/CONSEJO (DETECTAR CAPACIDADES/INTERESES) • DESCRIPCIÓN/INFORMACIÓN (DESCRIBIR RENDIMIENTO/INFORMACIÓN OPINIÓN PÚBLICA)		
ESPECIFICAR CARACTERÍSTICAS TEST	CONTENIDO UNA VEZ DEFINIDO EL CONSTRUCTO Y SU DOMINIO DE CONDUCTAS	PARA MEJOR DEFINICIÓN DEL CONSTRUCTO Y DOMINIO DE CONDUCTAS: • ANÁLISIS DEL CONTENIDO DE CONSTRUCTO (PREGUNTAS ABIERTAS) • REVISAR INVESTIGACIONES PUBLICADAS E INSTRUMENTOS DE MEDICIÓN 3 • ANÁLISIS DE TAREAS (CONDUCTAS CRÍTICAS) • OBSERVACIÓN DIRECTA POR PARTE DEL CONSTRUCTOR DEL TEST • UTILIZAR EL JUICIO DE EXPERTOS • REVISAR OBJETIVOS INSTRUCCIONALES O DE PROGRAMAS DE INTERVENCIÓN		
		TEST DE RENDIMIENTO Y CONOCIMIENTOS y TRC: DOMINIO DE CONTENIDOS MUY ESPECÍFICO. TABLA DE DOBLE ENTRADA - COLUMNAS: ÁREAS DE CONTENIDO (% ÍTEMS) - FILAS: PROCESOS U OPERACIONES (CATEGORIZACIÓN JERÁRQUICA, TAXONOMÍA DE BENJAMIN BLOOM : CONOCIMIENTO, COMPRENSIÓN, APLICACIÓN, ANÁLISIS, SÍNTESIS, EVALUACIÓN) O GRADO DE DIFICULTAD . TEST DE APTITUDES MENOS ESPECÍFICOS: PODEMOS DEFINIR TIPOS DE ÍTEMS (HABILIDAD VERBAL: ANALOGÍAS, SINÓNIMOS, ANTÓNIMOS, ÍTEMS DE FRASES INCOMPLETAS...)		
	FORMATO DE LOS ÍTEMS V. TIPO COGNITIVO: TEST OBJETIVOS V. TIPO NO COGNITIVO	ÍTEMS DE ELECCIÓN RESPUESTA CERRADA. EL SUJETO SELECCIONA SU RESPUESTA.	1. DOS ALTERNATIVAS V. TIPO COGNITIVO FORMATO: ENUNCIADO Y DOS ALTERNATIVAS: V/F, S/N, CORRECTO/INCORRECTO. "LOS CONSTRUCTOS SON VARIABLES OBSERVABLES DIRECTAMENTE V () F ()"	VENTAJAS: RÁPIDO Y FÁCIL DE USAR INCONVENIENTES: - 50% ACIERTO POR AZAR. - PARA LOS TEST T. NO COGNITIVO POCA VARIEDAD PARA REFLEJAR POSICION.
			2. ELECCIÓN MÚLTIPLE V. TIPO COGNITIVO FORMATO: ENUNCIADO Y VARIAS ALTERNATIVAS (CORRECTA Y DISTRACTORES) DE 3 A 5 PARA REDUCIR ACIERTO POR AZAR. ELECCIÓN DE LA CORRECTA (NO HAY AMBIGÜEDAD) O LA MÁS COMPLETA (PROCESOS MENTALES COMPLEJOS) ENUNCIADO: FORMA INTERROGATIVA (LA MÁS DIRECTA Y RECOMENDABLE), ENUNCIATIVA, FRASE TRUNCADA O INCOMPLETA (TEST EDUCATIVOS).	VENTAJAS: - FÁCILES DE ADMINISTRAR, CORREGIR Y PUNTUAR (LECTORES ÓPTICOS Y ORDENADOR) - REDUCE ACIERTO POR AZAR AL 33-20%. INCONVENIENTES: MÁS DIFÍCILES DE CONSTRUIR. "UNA DE LAS RECOMENDACIONES EN LA REDACCIÓN DE ÍTEMS ES: A) QUE EXPRESE MÁS DE UNA IDEA EN EL ENUNCIADO; B) LA UTILIZACIÓN DE DOBLES NEGATIVOS; C) QUE DEN LUGAR A RESPUESTAS INSEGADAS"
			3. EMPAREJAMIENTO V. TIPO COGNITIVO (CONOCIMIENTOS) FORMATO: EMPAREJAR ELEMENTOS DE DOS COLUMNAS.	VARIABLE NOMINAL COEFICIENTE PHI () VARIABLE ORDINAL COEFICIENTE C () VARIABLE CUANTITATIVA COEFICIENTE DE SPEARMAN () VARIABLE DICOTÓMICA CORRELACIÓN PEARSON ()
			4. CLOZE O INCOMPLETO V. TIPO COGNITIVO FORMATO: UN PÁRRAFO O UNA FRASE EN LA QUE FALTAN ALGUNAS PALABRAS Y LAS PALABRAS QUE FALTAN. UBICAR LA QUE CORRESPONDA EN CADA ESPACIO. "EL MÉTODO DE NEDLSKY, ES UN MÉTODO.....PARA ESTIMAR EL PUNTO DE CORTE EN LOS TESTS REFERIDOS AL CRITERIO. A) VALORATIVO; B) DE COMPROMISO; C) COMBINADO"	

			5. ESCALAS DE CLASIFICACIÓN (RATING SCALES) V. TIPO NO COGNITIVO (no R correcta o incorrecta) FORMATO: ENUNCIADO Y <u>DISTINTAS ALTERNATIVAS</u> PERTENECIENTES A <u>CATEGORÍAS INTERDEPENDIENTES Y ORDENADAS GRADUALMENTE</u> (EL TIPO MÁS UTILIZADO ES EL DE 5. CUANDO HAY MÁS DE 7 ES CASI IMPOSIBLE DIFERENCIAR SUS SIGNIFICADOS) ELEGIR LA QUE MEJOR REFLEJE POSTURA O ACTITUD PERSONAL. <u>ESCALAS VALORATIVAS.</u> ETIQUETAS LINGÜÍSTICAS: ACUERDO (TOTALMENTE DE ACUERDO, TOTALMENTE EN DESACUERDO), FRECUENCIA (SIEMPRE, NUNCA), CANTIDAD (MUCHO, NADA), SATISFACCIÓN (COMPLETAMENTE S, COMPLETAMENTE INSATISFECHO) VALORACIÓN (EXCELENTE, MUY MALA)		VENTAJAS: - REFLEJAN POSTURA DE MANERA MÁS PRECISA . INCONVENIENTES: - EL SIGNIFICADO NO ES EL MISMO PARA TODOS LOS SUJETOS. - SESGOS: TENDENCIA A ELECCIÓN DE <u>OPCIONES EXTREMAS O CENTRAL.</u> “EL TABACO DEBERÍA PROHIBIRSE EN TODOS LOS SITIOS PÚBLICOS: A) TOTALMENTE DE ACUERDO; B) DE ACUERDO; C) ME ES INDIFERENTE; D) EN DESACUERDO; E) TOTALMENTE EN DESACUERDO”																																																
			6. LISTADOS (CHESKLISTS) V. TIPO NO COGNITIVO (no R correcta o incorrecta) FORMATO: ENUNCIADO Y <u>ALTERNATIVAS INDEPENDIENTES</u> ENTRE SÍ, NO ESTÁN ORDENADAS. EL NÚMERO DE ALTERNATIVAS ES BASTANTE GRANDE (LISTADO) Y ES POSIBLE ELEGIR VARIAS OPCIONES. ELEGIR LA QUE MEJOR REFLEJE OPINIÓN. <u>ESCALAS VALORATIVAS.</u>		<table><tr><td></td><td></td><td>Si</td><td>No</td><td>NC</td></tr><tr><td rowspan="9">¿Cuáles de los siguientes bienes poseen en su hogar?</td><td>TV color</td><td></td><td></td><td></td></tr><tr><td>Heladera o refrigerador</td><td></td><td></td><td></td></tr><tr><td>Vídeo grabador / reproductor</td><td></td><td></td><td></td></tr><tr><td>Aire acondicionado</td><td></td><td></td><td></td></tr><tr><td>Teléfono</td><td></td><td></td><td></td></tr><tr><td>Tarjeta de crédito</td><td></td><td></td><td></td></tr><tr><td>Computador/ a</td><td></td><td></td><td></td></tr><tr><td>Lavadora o lavadero</td><td></td><td></td><td></td></tr><tr><td>Agua potable</td><td></td><td></td><td></td></tr><tr><td>Agua caliente</td><td></td><td></td><td></td></tr><tr><td>Alcantarillado o cloacas</td><td></td><td></td><td></td></tr></table>			Si	No	NC	¿Cuáles de los siguientes bienes poseen en su hogar?	TV color				Heladera o refrigerador				Vídeo grabador / reproductor				Aire acondicionado				Teléfono				Tarjeta de crédito				Computador/ a				Lavadora o lavadero				Agua potable				Agua caliente				Alcantarillado o cloacas	
			Si	No	NC																																																
	¿Cuáles de los siguientes bienes poseen en su hogar?	TV color																																																			
Heladera o refrigerador																																																					
Vídeo grabador / reproductor																																																					
Aire acondicionado																																																					
Teléfono																																																					
Tarjeta de crédito																																																					
Computador/ a																																																					
Lavadora o lavadero																																																					
Agua potable																																																					
Agua caliente																																																					
Alcantarillado o cloacas																																																					
ÍTEMS DE CONSTRUCCIÓN RESPUESTA ABIERTA. EL SUJETO ELABORA SU RESPUESTA.	1. ÍTEMS DE RESPUESTA CORTA FORMATO: <u>“CONSTRUIR/ELEGIR” UNA PALABRA O FRASE</u> PARA COMPLETAR UN ENUNCIADO: “EL NOMBRE DEL PRESIDENTE DEL GOBIERNO ESPAÑOL ES.....”		VENTAJAS: - SON MÁS FÁCILES DE CONSTRUIR . - PERMITEN HACER UN <u>ANÁLISIS CUALITATIVO</u> , DE ASPECTOS COMO: ORIGINALIDAD, FORMA DE REDACTAR... ES DECIR, HABILIDADES COGNITIVAS DE ORDEN SUPERIOR. INCONVENIENTES: - LAS RESPUESTAS SON MÁS DIFÍCILES DE ANALIZAR Y VALORAR . REQUIERE CODIFICACIÓN PREVIA EN CATEGORÍAS A SUJETOS CON R SIMILARES - JUICIOS SUBJETIVOS.																																																		
	2. ÍTEMS DE RESPUESTA EXTENSA O ENSAYO FORMATO: <u>DESARROLLAR</u> UN TEMA. “ESCRIBE UNA DESCRIPCIÓN DE TI MISMO/A COMO SI FUERAS EL/LA PROTAGONISTA DE UNA OBRA DE TEATRO. ESCRÍBELO COMO PODRÍA ESCRIBIRLO UN AMIGO QUE TE CONOCE MEJOR QUE NINGUNA OTRA PERSONA. ASEGÚRATE DE ESCRIBIRLO EN TERCERA PERSONA”																																																				
LONGITUD DEL TEST <i>PRUEBA PILOTO</i> MAYOR NÚMERO DE ÍTEMS QUE VERSIÓN FINAL	<u>POBLACIÓN A LA QUE VA DIRIGIDO</u>	EJ. POBLACIÓN INFANTIL VS ADULTA (≠ TIEMPO PROCESAMIENTO Y EJECUCIÓN, CAP. DE ATENCIÓN Y MOTIVACIÓN)																																																			
	<u>TIEMPO DE EJECUCIÓN DISPONIBLE</u>	NORMA GENERAL: <u>TIEMPO SUFICIENTE PARA INTENTAR CONTESTAR TODOS LOS ÍTEMS. 1 MIN ÍTEM</u> (EXCEPTO SI MEDIMOS RAPIDEZ DE RESPUESTA)																																																			
	<u>OBJETIVOS DEL TESTS</u>	ÁREA DE CONOCIMIENTO CONCRETA: <u>ÍTEMS ESPECÍFICOS/TEST CORTO</u> VARIAS ÁREAS DE CONTENIDO: <u>MÁS VARIADO Y MAYOR Nº ÍTEMS</u>																																																			
CARACTERÍSTICAS PSICOMÉTRICAS DE LOS ÍTEMS	NIVEL DE DIFICULTAD	<u>PROBABILIDAD DE RESPONDERLE CORRECTAMENTE.</u> <u>TESTS DE VELOCIDAD:</u> <u>ÍTEMS MUY FÁCILES, EL TIEMPO ES EL QUE DISCRIMINA</u> (SIN LIMITACIÓN CASI TODOS PODRÍAN RESOLVERLOS) EN V. COGNITIVAS: VEL. CÁLCULO. <u>TESTS DE EJECUCIÓN MÁXIMA O DE POTENCIA</u> (RAVEN; DAT) <u>DIFERENTES GRADOS DE DIFICULTAD (MUY FÁCIL, MUY DIFÍCIL) ORDENADOS.</u> <u>DEBEN TENER TIEMPO PARA CONTESTAR A TODOS.</u> CUANDO NO SE CONTESTA ES PORQUE NO SE CONOCE LA RESPUESTA. <u>TESTS DE EJECUCIÓN TÍPICA</u> (16 –PF, MMPJ): NO TIENE SENTIDO HABLAR DE DIFICULTAD																																																			
	GRADO DE HOMOGENEIDAD	<u>EN F. DE QUE MIDA LO MISMO QUE EL RESTO.</u> SI EL CONSTRUCTO ES <u>UNIDIMENSIONAL</u> LOS ÍTEMS HAN DE SER MÁS <u>HOMOGÉNEOS</u> QUE SI ES MULTIDIMENSIONAL.																																																			
	NIVEL DE DISCRIMINACIÓN	<u>EN F. DE QUE SIRVA PARA DIFERENCIAR SUJETOS DE PUNTUACIONES EXTREMAS EN EL TEST.</u> - PARA LA <u>POBLACIÓN GENERAL</u> LOS ÍTEMS DEBEN DISCRIMINAR ENTRE LOS DISTINTOS NIVELES QUE PRESENTAN LOS SUJETOS: MAYOR PORCENTAJE DE <u>ÍTEMS DE DIFICULTAD MEDIA.</u> - PARA DETECTAR A LOS <u>MÁS BRILLANTES</u> Y DISCRIMINAR ENTRE ELLOS, DEBEN SER <u>ÍTEMS DIFÍCILES Y MUY DIFÍCILES.</u> - PARA DISCRIMINAR ENTRE LOS <u>MENOS CAPACES, FÁCILES Y MUY FÁCILES.</u>																																																			

REDACCIÓN DE LOS ÍTEMS ALTO GRADO DE PRECISIÓN EN EL USO DEL LENGUAJE Y CONOCIMIENTO PROFUNDO DEL CONTENIDO	PARA LA CONSTRUCCIÓN DE UN BUEN TEST, DEBEMOS CONSTRUIR BUENOS ÍTEMS. CONSIDERACIONES:		3. MINIMIZAR LOS ERRORES DE MEDIDA
	1. <u>ALTO GRADO DE CONGRUENCIA ENTRE ÍTEM Y CONSTRUCTO (VALIDEZ DE CONSTRUCTO)</u>		4. <u>FORMATO ADECUADO</u> A LOS OBJETIVOS DEL TEST
	2. CONSTRUCTOS CLARAMENTE DEFINIDOS		5. <u>CARACTERÍSTICAS PSICOMÉTRICAS</u> MÁS ADECUADAS EN CADA CASO.
			6. <u>BIEN REDACTADOS</u>
			7. <u>SATISFACER LAS CONSIDERACIONES LEGALES Y TÉCNICAS PERTINENTES (EVITAR LOS PLAGIOS)</u>
RECOMENDACIONES GENERALES PARA LA REDACCIÓN		• <u>EVITAR LA AMBIGÜEDAD</u> DE LOS ENUNCIADOS: <u>ENUNCIADOS CORTOS Y DIRECTOS</u>. CLAROS Y PRECISOS. <u>SIN TÉRMINOS INTERPRETABLES</u> (PATRIOTISMO, RELIGIOSIDAD...) • <u>EVITAR</u> ENUNCIADOS QUE PROVOQUEN <u>RESPUESTAS SESGADAS</u> (QUE LAS ELIJAN INDEPENDIENTEMENTE DE SU OPINIÓN: <u>DESEABILIDAD SOCIAL</u>) • EXPRESAR UNA <u>ÚNICA IDEA</u> EN EL ENUNCIADO: <u>EVITAR DOBLES PREGUNTAS</u>. • <u>EVITAR LAS DOBLES NEGACIONES</u> EN LOS ENUNCIADOS: <u>TAMPOCO ABUSAR DE LOS ENUNCIADOS NEGATIVOS</u>.	
ÍTEMS DE ELECCIÓN		DOS ALTERNATIVAS: - ESTAR SEGURO DE QUE EFECTIVAMENTE ES VERDADERO O FALSO. - NO UTILIZAR FRASES QUE SEAN UNIVERSALMENTE VR O FALSAS. - EVITAR INDUCIR LA RESPUESTA CORRECTA (SIEMPRE, TODO, NADA, NUNCA, EXCLUSIVAMENTE: FALSO; A VECES, EN GENERAL, APENAS...: VR) - SITUAR DE FORMA ALEATORIA LOS VR PARA EVITAR PATRONES DE RESPUESTA. ELECCIÓN MÚLTIPLE: - CLARIDAD - INCLUIR LA MAYOR PARTE DEL TEXTO EN EL ENUNCIADO - INCLUIR LAS DISTINTAS OPCIONES DE R EN EL ENUNCIADO - DISTRACTORES PLAUSIBLES - EVITAR: <i>NINGUNA DE LAS ANTERIORES, TODAS LAS ANTERIORES.</i> - SOLO UNA CORRECTA O COMPLETA - ALTERNATIVAS CON LONGITUD Y CONSTRUCCIÓN GRAMATICAL PARECIDA ALEATORIZAR UBICACIÓN DE LA CORRECTA - ALTERNATIVAS IGUAL DE ATRACTIVAS PARA UNA PERSONA NO INFORMADA. - ALTERNATIVAS CON CONCORDANCIA GRAMATICAL CON EL ENUNCIADO EMPAJEAMIENTO: - PREMISAS Y OPCIONES DE RESPUESTA HOMOGÉNEAS - FORMATO ADECUADO Y ENUNCIADO CLARO FORMATO CLOZE O INCOMPLETO - TANTOS ESPACIOS EN BLANCO COMO ALTERNATIVAS ESCALAS DE CLASIFICACIÓN: - EVITAR EXPRESIONES COLOQUIALES - MISMO Nº DE ÍTEMS FORMULADOS DE MANERA POSITIVA Y NEGATIVA - ASIGNAR ETIQUETAS LINGÜÍSTICAS: CATEGORÍA CENTRAL QUE REPRESENTA PUNTO MEDIO O NEUTRAL. LISTADOS: FÁCILES DE CONSTRUIR.	
ÍTEMS DE CONSTRUCCION		DE RESPUESTA CORTA: - ASEGURARSE QUE PUEDE SER CONTESTADO CON UNA ÚNICA FRASE O PALABRA Y UNA ÚNICA RESPUESTA. - ESPACIOS EN BLANCO DE LA MISMA LONGITUD - EVITAR DAR PISTAS O CLAVES (ARTÍCULOS, ETC) - INDICAR GRADO DE PRECISIÓN REQUERIDO. - NO DETERMINANTES ESPECÍFICOS (TODO/NADA) Y AMBIGUOS (FRECUENT/, ALGUNAS VCS) DE RESPUESTA EXTENSA O ENSAYO: - ASEGURARSE BUEN ENFOQUE. DEFINIR CLARAMENTE: COMPARE, CONTRASTE... PARA EVITAR VAGUEDADES... - NO PERMITIR ELEGIR ENTRE VARIAS PREGUNTAS: PARA PODER COMPARAR RENDIMIENTO - DECIDIR DE ANTEMANO CÓMO SE VA A PUNTUAR: REGLAS COMPLETAS Y EXPLÍCITAS. - LAS PREGUNTAS REFERIDAS A CUESTIONES CONTROVERTIDAS EVALUADAS REFERIDAS A LA EVIDENCIA QUE PRESENTAN Y NO A SU POSICIÓN PERSONAL RESPECTO AL TEMA.	
SESGOS DE RESPUESTA TENDENCIAS		APARECEN EN TEST QUE MIDEN ASPECTOS ORÉCTICOS Y AFECTIVOS: • <u>AQUIESCENCIA</u> (TENDENCIA A RESPONDER SISTEMÁTICA/ QUE SE ESTÁ DE ACUERDO/EN DESACUERDO CON EL ÍTEM CON INDEPENDENCIA DEL CONTENIDO) • <u>DESEABILIDAD SOCIAL</u> • <u>INDECISIÓN</u> (ALTERNATIVA NEUTRA) • <u>RESPUESTA EXTREMA</u>	
REVISIÓN CRÍTICA GRUPO DE EXPERTOS		<u>REVISIÓN DE CONTENIDO Y ESTILO</u> POR UN GRUPO DE PERSONAS <u>QUE NO HAYAN PARTICIPADO EN SU ELABORACIÓN</u> . UNA VEZ REVISADOS LOS ÍTEMS ELIMINADOS (O CORREGIDOS) AQUELLOS QUE NO FUERAN CONSIDERADOS IDÓNEOS, PODEMOS CONSTRUIR LA VERSIÓN PRELIMINAR DEL TEST: <u>LA PRUEBA PILOTO</u> .	
CONFECCIÓN DE LA PRUEBA PILOTO		INSTRUCCIONES ADMINISTRACIÓN	- EVITAR LENGUAJE AMPULOSO Y AMENAZANTE - <u>EN LOS TEST DE EJECUCIÓN MÁXIMA</u> EXPLICITAR QUE HAY ÍTEMS DE DIFICULTAD VARIABLE PARA REDUCIR LA ANSIEDAD. - <u>EN LOS TEST DE VELOCIDAD</u> EXPLICITAR QUE ES MUY POCO TIEMPO Y QUE SOLO MUY POCOS LLEGARÁN A COMPLETAR LA PRUEBA. - <u>PROPORCIONAR ÍTEMS COMO EJEMPLO</u>. - INFORMAR SOBRE CÓMO DISTRIBUIR TIEMPO Y QUÉ HACER CUANDO NO SABEMOS LA RESPUESTA. CUANDO HAY TIEMPO LÍMITE PEDIR RAPIDEZ. - <u>ANIMAR A RESPONDER A TODAS LAS PREGUNTAS</u> PARA FAVORECER EL RENDIMIENTO. - EXPLICITAR LA FORMA DE RESPONDER EN HOJAS DE LECTURA ÓPTICA.
		FORMATO DE PRESENTACIÓN	- CLARO Y LEGIBLE. SOLICITAR DATOS O CLAVE DE IDENTIFICACIÓN. - <u>ORGANIZAR Y ELABORAR ÍTEMS:</u> EN VARIABLES COGNITIVAS <u>ORDENADOS SEGÚN NIVEL DE DIFICULTAD Y POR TEMAS</u> (DESMOTIVACIÓN) EN VARIABLES NO COGNITIVAS <u>LAS EMBARAZOSAS AL FINAL</u>. <u>AGRUPADOS POR FORMATOS</u> SI HAY VARIOS.
		REGISTRO DE RESPUESTAS	- EN LA MISMA HOJA O CUADERNILLO DEL TEST - EN HOJA APARTE. LECTORA ÓPTICA. REAPROVECHAMOS CUADERNILLOS.

APLICACIÓN DE LA PRUEBA PILOTO	- <u>MUESTRA REPRESENTATIVA</u> DE LA POBLACIÓN A LA QUE VA DIRIGIDO EL TEST - DECIDIR FORMA DE ADMISTRACIÓN: • <u>INDIVIDUAL</u> VS <u>COLECTIVA</u>: SIEMPRE QUE SE PUEDA COLECTIVA. • <u>ORAL</u> (PERSONAL/TELFÓNICA): NIÑOS PEQUEÑOS/ENCUESTAS. • <u>PAPEL Y LÁPIZ</u> VS <u>ORDENADOR</u> (MENOS COSTE DE TIEMPO, MAYOR ESTANDARIZACIÓN) • <u>POR CORREO</u>: SE ENVÍA EL TEST, UNA CARTA EXPLICATIVA Y UN SOBRE TIMBRADO PARA REMITIR. SE HACE EN ESTUDIOS DE OPINIÓN Y AQUELLOS QUE REQUIERAN CONSULTA DE INFORMACIÓN PARA SU REALIZACIÓN. <u>VENTAJA</u> SE PUEDEN ENVIAR UN GRAN NÚMERO DE PRUEBAS. <u>INCONVENIENTE</u> ALTA TASA DE PERSONAS QUE NO RESPONDEN (50%). FALTA DE SEGURIDAD DE LA PERSONA QUE RESPONDE.			
CORRECCIÓN DE LA PRUEBA PILOTO Y ASIGNACIÓN DE PUNTUACIONES A LOS SUJETOS	ÍTEMS DE ELECCIÓN	<u>PRUEBAS COGNITIVAS</u> PUNTUACIÓN $X = \sum X_i$, <u>PARA CORREGIR EL AZAR</u>:	PUNTUACIÓN PENALIZANDO ERRORES	<u>PUNTUACIÓN CORREGIDA</u>: $X_c = A - A_a = A - \frac{E}{(k-1)}$ PROBABILIDAD DE ACIERTO POR AZAR: $P(A_a) = \frac{1}{k}$; ACIERTOS POR AZAR EN TEST: $A_a = R_a \frac{1}{k}$, PROBABILIDAD DE ERROR: $P(E) = 1 - \frac{1}{k}$; ERRORES EN TEST: $E = R_a \left(1 - \frac{1}{k}\right)$ DESPEJANDO R_a Y SUSTITUYÉNDOLA EN LA DE ACIERTOS POR AZAR EN TEST: $A_a = E \frac{k}{k-1} \left(\frac{1}{k}\right) = \frac{E}{k-1}$ SIENDO K EL ° DE ALTERNATIVAS Y R_a RESPUESTAS ALEATORIAS
		*CUANDO UN TEST ESTÁ FORMADO POR ÍTEMS CON <u>DISTINTO NÚMERO DE ALTERNATIVAS</u> LA <u>CORRECCIÓN DE AZAR SE HARÁ POR PARTES</u>.	PUNTUACIÓN BONIFICANDO OMISIONES	AÑADIMOS ACIERTOS QUE HUBIERA TENIDO SI EN VEZ DE NO CONTESTAR HUBIERA CONTESTADO AL AZAR <u>PUNTUACIÓN CORREGIDA</u>: $X_c = A + A_a = A + \frac{O}{k}$; $R_a = O$ PUNTUACIÓN SOBREVALORADA. NO CORRESPONDE AL VERDADERO NIVEL DEL SUJETO. ES MEJOR PENALIZAR EL ERROR. SIENDO K EL ° DE ALTERNATIVAS, $R_a = O$ RESPUESTAS ALEATORIAS Y O OMISIONES
		<u>PRUEBAS NO COGNITIVAS</u> PUNTUACIÓN SUMANDO LOS VALORES NUMÉRICOS ASIGNADOS A LA CATEGORÍAS ELEGIDAS.	<u>ESCALAMIENTO PREVIO DE LOS ÍTEMS EN FUNCIÓN DEL GRADO DE ATRIBUTO</u> (VALOR NUMÉRICO DISTINTO A CADA ALTERNATIVA DE RESPUESTA O CATEGORÍA) <u>TENER MUY CLARA LA DIRECCIÓN DEL CONTINUO</u> DE LA VARIABLES Y ASIGNAR A CADA EXTREMO EL VALOR NUMÉRICO MÁS ALTO Y MÁS BAJO RESPECTIVAMENTE.	
	ÍTEMS DE CONSTRUCCIÓN <u>CORRECCIÓN Y CONTROL DE LA SUBJETIVIDAD</u>	EN LOS ÍTEMS DE RESPUESTA CORTA ES MÁS FÁCIL ASIGNAR UNA PUNTUACIÓN. <u>SE COMPLICA A MEDIDA QUE LA RESPUESTA ES MÁS ABIERTA Y EXTENSA</u>. CORRECCIÓN BASTANTE SUBJETIVA.		
		MÉTODO DE LA PUNTUACIÓN <u>ANALÍTICA</u>		- DEFINIR DE FORMA INEQUÍVOCA Y AISLAR LAS DIMENSIONES MÁS IMPORTANTES - DEFINIR LA FORMA DE EVALUARLAS: QUÉ ES CORRECTO EN CADA DIMENSIÓN. QUÉ RESPUESTAS SON NECESARIAS PARA CONSIDERARLAS COMPLETAS. - <u>PUEDEN SER CORREGIDAS POR PERSONAS NO EXPERTAS</u> EN EL TEMA.
		MÉTODO DE LA PUNTUACIÓN <u>HOLÍSTICA</u>		- LA PUNTUACIÓN EXPRESA LA CALIDAD GLOBAL DE LA RESPUESTA TOMANDO DISTINTOS VALORES DENTRO DE LOS LÍMITES ESTABLECIDOS. - REQUIERE SER HECHA <u>POR EXPERTOS</u> (PREVIAMENTE HABIENDO LLEGADO A UN ACUERDO)

TEMA 3. TÉCNICAS PARA LA CONSTRUCCIÓN DE ESCALAS DE ACTITUDES

PROCEDIMIENTOS

CONSTRUCCIÓN DE ESCALAS, CUESTIONARIOS E INVENTARIOS PARA LA MEDICIÓN DE VARIABLES NO COGNITIVAS : PERSONALIDAD, ACTITUDES, INTERESES, VALORES, OPINIONES, ETC.		
- EL MODELO ESCALAR DE THURSTONE	LA LEY DEL JUICIO COMPARATIVO → LA LEY DEL JUICIO CATEGÓRICO {	- COMPARACIONES BINARIAS - ORDENACIÓN POR RANGOS - INTERVALOS APARENTEMENTE IGUALES - INTERVALOS SUCESIVOS
- EL DIFERENCIAL SEMÁNTICO DE OSGOOD		
- EL ESCALOGRAMA DE GUTTMAN		
- LA TÉCNICA DE LIKERT		

DISTINCIÓN ENTRE INTERESES, VALORES Y ACTITUDES SEGÚN NUNNALLY

INTERESES : PREFERENCIAS POR ACTIVIDADES PARTICULARES <i>"PREFIERO TRABAJAR AL AIRE LIBRE QUE EN UN DESPACHO"</i>	} ESCALAS DE INTERESES Y VALORES ÍTEMS REFERIDOS A NUMEROSAS ACTIVIDADES (ESPECÍFICAS PARA MEDIR INTERESES Y AMPLIAS PARA MEDIR VALORES)
VALORES : PREFERENCIAS SOBRE OBJETIVOS Y FORMAS DE VIDA MÁS QUE SOBRE ACTIVIDADES CONCRETAS <i>"CONSIDERO MÁS IMPORTANTE QUE ME RESPETEN A QUE ME QUIERAN"</i>	
ACTITUDES : SENTIMIENTOS ACERCA DE UN DETERMINADO OBJETO SOCIAL (FÍSICOS, TIPOS DE PERSONAS, INSTITUCIONES SOCIALES, POLÍTICAS,...) <i>"TODAS LAS GUARDERÍAS PÚBLICAS DEBERÍAN SER GRATUITAS"</i>	} ESCALAS DE ACTITUDES ÍTEMS REFERIDOS AL MISMO OBJETO SOCIAL (MISMA VARIABLE)

1. EL MODELO ESCALAR DE THURSTONE

- EL OBJETIVO ES ESCALAR LOS ESTÍMULOS A LO LARGO DE UN CONTINUO PSICOLÓGICO, ASIGNANDO UN VALOR EN LA ESCALA A CADA UNO DE ELLOS. - SE BASA EN LA VARIABILIDAD PERCEPTUAL DE LOS SUJETOS (DISTINTOS SUJETOS, DISTINTOS MOMENTOS) Y EN LA LIMITACIÓN QUE TIENEN LOS SUJETOS PARA PERCIBIR DIFERENCIA DE MAGNITUD ENTRE DOS ESTÍMULOS (CUANTO MAYOR SEA LA DIFERENCIA MÁS FÁCIL DIFERENCIARLOS, ORDENARLOS Y ASIGNARLES VALOR) - LA DIFERENCIA FUNDAMENTAL CON LAS ESCALAS ELABORADAS POR OTROS PROCEDIMIENTOS ES QUE HAY QUE AÑADIR UNA NUEVA FASE, LA PRUEBA DE JUECES, EN LA QUE SE ASIGNAN VALORES ESCALARES (PUNTUACIONES) A CADA UNO DE LOS ÍTEMS (ESTÍMULOS)	
SUPUESTOS BÁSICOS	EXISTE ESE CONTINUO PSICOLÓGICO O SUBJETIVO A LO LARGO DEL CUAL VARÍA EL ATRIBUTO. - CADA UNO DE LOS ESTÍMULOS ORIGINA EN EL SUJETO UN PROCESO DISCRIMINANTE (SUBJETIVO) A TRAVÉS DEL CUAL LES ASIGNARÁ UN VALOR TAMBIÉN SUBJETIVO EN EL CONTINUO PSICOLÓGICO. EL VALOR SUBJETIVO ASIGNADO POR EL SUJETO EN CADA OCASIÓN AL ESTÍMULO PUEDE VARIAR. SI PRESENTAMOS EL ESTÍMULO UN NÚMERO ALTO DE VECES LA DISTRIBUCIÓN DE LOS VALORES SUBJETIVOS SE AJUSTAN A LA DISTRIBUCIÓN NORMAL → DISTRIBUCIÓN DISCRIMINANTE. - LA MEDIA (VALOR ESCALAR DEL ESTÍMULO) ES EL VALOR DEL ESTÍMULO EN EL CONTINUO PSICOLÓGICO. - LA DESVIACIÓN TÍPICA (DISPERSIÓN DISCRIMINANTE) NOS DA IDEA DE LA AMBIGÜEDAD SUSCITADA POR EL ESTÍMULO EN EL SUJETO. - SI SE PRESENTAN A UN MISMO SUJETO VARIOS ESTÍMULOS, CADA UNO DARÁ UNA DISTRIBUCIÓN DISCRIMINANTE DISTINTA DEFINIDA POR SU MEDIA Y POR SU DESVIACIÓN. - EL MODELO SE CUMPLE TANTO SI ES UN SOLO SUJETO EL QUE EMITE NUMEROSOS JUICIOS SOBRE CADA UNO DE LOS ESTÍMULOS, COMO SI ES UNA MUESTRA GRANDE DE SUJETOS (JUECES O EXPERTOS) QUE EMITEN UN SOLO JUICIO SOBRE CADA ESTÍMULO → PRUEBA DE JUECES. - LOS JUECES EMITEN JUICIOS DE HECHO, NO JUICIOS DE VALOR. NO SE PRETENDE QUE EL JUEZ VALORE LA ACTITUD, SINO QUE EVALÚE SI EL ÍTEM ES CAPAZ, Y EN QUÉ GRADO, DE MEDIRLA.

LEY DEL JUICIO COMPARATIVO

- MÉTODO EXPERIMENTAL PARA LA OBTENCIÓN EMPÍRICA DE LOS DATOS: MÉTODO DE COMPARACIONES BINARIAS.	
	COMBINACIONES BINARIAS POSIBLES A PRESENTAR AL GRUPO DE JUECES: $C.B. = \frac{n(n-1)}{2}$, siendo n: n° de ítems
- LA TAREA DE LOS SUJETOS ES COMPARAR DIRECTAMENTE CADA UNO DE LOS ESTÍMULOS QUE SE LE PRESENTAN CON TODOS LOS DEMÁS.	
- LOS JUECES DEBEN INDICAR CUAL ES EL ÍTEM DE CADA PAR QUE, A SU JUICIO, ES DOMINANTE EN LA DIRECCIÓN DEL ATRIBUTO QUE SE ESTÁ MIDIENDO.	

MATRICES

- ENTRE CADA PAR DE ESTÍMULOS SE PRODUCE UNA DIFERENCIA DISCRIMINANTE, PUDIENDO OBTENER TRAS LA PRUEBA DE JUECES LA DISTRIBUCIÓN DE FRECUENCIAS DE CADA ESTÍMULO. ESTA DISTRIBUCIÓN SERÁ NORMAL, CON MEDIA LA DIFERENCIA DE LOS VALORES ESCALARES DE LOS ESTÍMULOS Y SU DESVIACIÓN TÍPICA.
- LOS RESULTADOS DE LOS JUICIOS EMITIDOS SE ORDENAN EN UNA SERIE DE MATRICES DE FRECUENCIAS, DE PROPORCIONES Y DE PUNTUACIONES TÍPICAS.
- LAS FILAS Y LAS COLUMNAS REPRESENTAN LOS DISTINTOS ÍTEMS QUE HAY QUE COMPARAR.

EJEMPLO. SE LES PIDE A LOS SUJETOS QUE JUZGUEN LA IMPORTANCIA DE CADA ELEMENTO (BIBLIO, TEATRO, CINE, GIMNASIO, AULA Y CAFETERÍA) EN RELACIÓN A SU EDUCACIÓN UNIVERSITARIA. $C.B. = \frac{6(6-1)}{2} = 15$, siendo $n = 6$ (n° de ítems)

Matriz de frecuencias

EE	Biblio	Teatro	Cine	Gim	Aula	Cafet
Biblio		70	65	45	40	80
Teatro	30		60	70	30	70
Cine	35	40		60	30	60
Gim	55	30	40		55	75
Aula	60	70	70	45		65
Cafet	20	30	40	25	35	
	200	240	275	245	190	350

Matriz de frecuencias ordenadas

EE	Cafet	Cine	Gim	Teatro	Biblio	Aula
Cafet		40	25	30	20	35
Cine	60		60	40	35	30
Gim	75	40		30	55	55
Teatro	70	60	70		30	30
Biblio	80	65	45	70		40
Aula	65	70	45	70	60	
	350	275	245	240	200	190

Matriz de proporciones ordenadas

estím ulos	Cafet	Cine	Gim	Teatro	Biblio	Aula
Cafet		0.40	0.25	0.30	0.20	0.35
Cine	0.60		0.60	0.40	0.35	0.30
Gim	0.75	0.40		0.30	0.55	0.55
Teatro	0.70	0.60	0.70		0.30	0.30
Biblio	0.80	0.65	0.45	0.70		0.40
Aula	0.65	0.70	0.45	0.70	0.60	
	3.50	2.75	2.45	2.40	2.00	1.90

Matriz de las puntuaciones típicas

estím ulos	Cafet	Cine	Gim	Teatro	Biblio	Aula
Cafet	0.00	-0.25	-0.67	-0.52	-0.84	-0.39
Cine	0.25	0.00	0.25	-0.25	-0.39	-0.52
Gim	0.67	-0.25	0.00	-0.52	0.13	0.13
Teatro	0.52	0.25	0.52	0.00	-0.52	-0.52
Biblio	0.84	0.39	-0.13	0.52	0.00	-0.25
Aula	0.39	0.52	-0.13	0.52	0.25	0.00
$\sum z_{ij}$	2.67	0.66	-0.16	-0.25	-1.37	-1.55
VE	0.45	0.11	-0.03	-0.04	-0.23	-0.26
Y	0.71	0.37	0.23	0.22	0.03	0.00

LAS CELDAS DE LA MATRIZ REPRESENTAN EL NUMERO DE JUECES QUE HAN CONSIDERADO QUE EL ESTÍMULO CORRESPONDIENTE A LA COLUMNA MUESTRA UN GRADO DE ACTITUD MAS FAVORABLE QUE EL REPRESENTADO EN LA FILA, ES DECIR EL NÚMERO DE VECES QUE EL OBJETO DE LA COLUMNA ES PREFERIDO SOBRE EL OBJETO DE LA FILA.

EN LA MATRIZ DE FRECUENCIAS EL MAYOR VALOR DEL SUMATORIO DE LAS COLUMNAS NOS MUESTRA LA ACTITUD MAS FAVORABLE. SE PODRÁ CONSTRUIR UNA ESCALA ORDINAL DE LOS ESTÍMULOS DESDE EL MÁS FAVORABLE AL MÁS DESFAVORABLE. LAS FILAS SE ORDENAN DE ACUERDO A LA ORDENACIÓN DE LAS COLUMNAS.

$VE = \frac{\sum z_{ij}}{n}$
PROPORCIÓN DE JUECES. SE OBTIENE DIVIDIENDO CADA UNO DE LOS ELEMENTOS DE LA MATRIZ ANTERIOR POR EL NÚMERO DE JUECES (N=100). LA SUMA DE ELEMENTOS SIMÉTRICOS ES IGUAL A LA UNIDAD.

LA MEDIA DE LAS PUNTUACIONES TÍPICAS ES LA MEJOR ESTIMACIÓN DE SU VALOR ESCALAR.

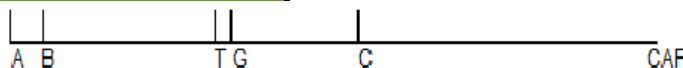
$$VE = \frac{\sum z_{ij}}{n}, \text{ siendo } \sum VE = 0$$

PUNTUACIONES TÍPICAS EN CADA CELDA. SE OBTIENEN BUSCANDO EN LAS TABLAS DE LA CURVA NORMAL LA PUNTUACIÓN QUE CORRESPONDE A CADA PROPORCIÓN.

COMO EN LA ESCALA EXISTEN VALORES NEGATIVOS HACEMOS UNA TRANSFORMACIÓN LINEAL DE LOS VALORES, TRASLADANDO EL ORIGEN AL VALOR ESCALAR MÁS BAJO: $Y = VE + |VE_{\text{mínimo}}|$

FINALMENTE OBTENEMOS LA ESCALA CON LOS VALORES ESTIMULARES. LA DISTANCIA ENTRE LOS DISTINTOS ÍTEMS VIENEN MARCADAS POR LAS DIFERENCIAS ENTRE SUS VALORES.

- ES SUBJETIVA, UNIDIMENSIONAL Y DE INTERVALOS.
- VA DESDE EL NIVEL DE ACTITUD MÁS DESFAVORABLE HASTA EL GRADO DE ACTITUD MÁS FAVORABLE.



LA LEY DEL JUICIO CATEGÓRICO

- DE LOS TRES PROCEDIMIENTOS PARA LA OBTENCIÓN EMPÍRICA DE LOS DATOS (ORDENACIÓN POR RANGOS, INTERVALOS SUCESIVOS E INTERVALOS APARENTEMENTE IGUALES) VEMOS EL ÚLTIMO POR SER EL MÁS COMÚNMENTE UTILIZADO: MÉTODO DE LOS INTERVALOS APARENTEMENTE IGUALES.
- ADEMÁS DE LOS SUPUESTOS GENERALES SE ASUME QUE EL CONTINUO PSICOLÓGICO PUEDE SER DIVIDIDO EN UNA SERIE DE CATEGORÍAS ORDENADAS.

LOS ESTÍMULOS SE EVALUARÁN SEGÚN UNA ESCALA IMAGINARIA DIVIDIDA EN CATEGORÍAS.

- DESDE LA ACTITUD MÁS NEGATIVA (CATEGORÍA MÍNIMA)
- HASTA LA MÁS POSITIVA (CATEGORÍA MÁXIMA)
- EN EL CENTRO SE ENCUENTRA EL PUNTO NEUTRAL, SEPARACIÓN ENTRE LA ACTITUD FAVORABLE Y LA DESFAVORABLE.

LOS LÍMITES REALES DE CADA CATEGORÍA SON:

- LÍMITE INFERIOR: $Li = (Li - 0,5)$
- LÍMITE SUPERIOR: $Ls = (Ls - 0,5)$

- LA TAREA DE LOS SUJETOS (JUECES MUESTRA GRANDE: $N > 100$) CONSISTE EN CLASIFICAR CADA UNO DE LOS ESTÍMULOS EN UNA DE LA CATEGORÍAS EN FUNCIÓN DEL GRADO Y LA DIRECCIÓN DEL ATRIBUTO QUE MANIFIESTEN. - SE OBTIENE PARA CADA ESTÍMULO UNA DISTRIBUCIÓN DE FRECUENCIAS DE LOS JUICIOS EMITIDOS Y SE ORDENAN EN FORMA DE MATRIZ DE FRECUENCIAS. - LAS COLUMNAS REPRESENTAN LAS CATEGORÍAS Y LAS FILAS LOS DISTINTOS ÍTEMS EVALUADOS. LAS CELDAS INDICAN EL NÚMERO DE VECES QUE CADA ESTÍMULO HA SIDO ASIGNADO A CADA CATEGORÍA. - EL VALOR ESCALAR DE LOS ESTÍMULOS ES LA MEDIANA DE LA DISTRIBUCIÓN DE FRECUENCIAS ORIGINADA POR CADA UNO DE ELLOS. PARA ELLO TENEMOS QUE HALLAR LAS FRECUENCIAS ACUMULADAS PARA CADA ÍTEM (f_d, f_b)		
	$VE = Md = l_i + \frac{I}{f_d} \left(\frac{N}{2} + f_b \right)$	I - AMPLITUD DEL INTERVALO ($I = 1$) f_b : Nº JUECES QUE CLASIFICÓ EL ELEMENTO EN CATEGORÍAS INFERIORES A LA DE LA MEDIANA f_d : Nº JUECES QUE CLASIFICÓ EL ELEMENTO DENTRO DE LA CATEGORÍA DE LA MEDIANA
FINALMENTE, PARA SELECCIONAR LOS ÍTEMS QUE FORMARÁN PARTE DE LA ESCALA SE COGEN AQUELLOS EN LOS QUE LOS JUECES <i>HAN MOSTRADO UN MAYOR ACUERDO, UNA MENOR AMBIGÜEDAD, O UNA MENOR DESVIACIÓN TÍPICA</i> .		
SE UTILIZA EL COEFICIENTE DE AMBIGÜEDAD (DISTANCIA ENTRE EL TERCER Y PRIMER CUARTIL) COMO MEDIDA DEL GRADO DE ACUERDO ENTRE LOS JUECES:	$C.A. = Q_3 - Q_1$ SI C.A. > 2 SE ELIMINA. EN CATEGORÍAS CENTRALES (NEUTRALES) EL C.A. > 3	$\left\{ \begin{array}{l} Q_3 = P_{75} = l_i + \frac{I}{f_d} \left(\frac{3N}{4} + f_b \right) \\ Q_1 = P_{25} = l_i + \frac{I}{f_d} \left(\frac{N}{4} + f_b \right) \end{array} \right.$ Q_3 PUNTUACIÓN POR DEBAJO DE LA CUAL QUEDAN EL 75% DE LOS JUECES Q_1 PUNTUACIÓN POR DEBAJO DE LA CUAL QUEDAN EL 25% DE LOS JUECES

- UNA VEZ QUE TENEMOS LA ESCALA (EN AMBOS JUICIOS) YA TENEMOS UN INSTRUMENTO DE MEDIDA Y PODEMOS HACER LA PRUEBA PILOTO CON UNA MUESTRA DE SUJETOS Y CONSTRUIR LA ESCALA DEFINITIVA. LOS SUJETOS EMITEN JUICIOS DE VALOR. EN FUNCIÓN DE SUS SENTIMIENTOS, POSICIÓN PERSONAL, ETC. RESPONDEN SI ESTÁN DE ACUERDO O NO CON EL ÍTEM. - LA PUNTUACIÓN EN LA ESCALA DE CADA SUJETO ES LA MEDIA DE LOS VALORES ESCALARES DE LOS ÍTEMS CON LOS QUE EL SUJETO ESTUVO DE ACUERDO. VENTAJA SOBRE OTRO TIPO DE ESCALAS ES QUE PERMITEN LA INTERPRETACIÓN DIRECTA DE LA ACTITUD DE UN SUJETO SIN NECESIDAD DE HACER REFERENCIA AL GRUPO O SIN RECURRIR A NORMAS GENERALES. EN PSICOLOGÍA Y SOCIOLOGÍA ESTO NO ES REALMENTE UNA VENTAJA. INCONVENIENTE: PUEDEN DEPENDER EN EXCESO DE LOS JUECES QUE SE UTILICEN.		
---	--	--

2. TÉCNICA DE LIKERT

UNO DE LOS MODELOS MÁS UTILIZADOS PARA LA MEDIDA DE LAS DIFERENCIAS INDIVIDUALES RESPECTO A LOS RASGOS PSICOLÓGICOS. - INTENTA DAR UNA SOLUCIÓN RAZONABLE A LOS ASPECTOS CUANTITATIVOS DEL ESTUDIO DE LAS ACTITUDES SOCIALES (SU ORIGEN SE SITÚA EN UN INVESTIGACIÓN HECHA POR GARDNER MURPHY, 1929) - LIKERT CONSIDERA QUE EL MÉTODO DE THURSTONE ES MUY LABORIOSO (PRUEBA DE JUECES, MÉTODOS ESTADÍSTICOS...) Y CREA UNA ESCALA FÁCIL DE CONSTRUIR, MUY FIABLE Y QUE PUEDE SER ADAPTADA. - ES UN CASO ESPECIAL DEL MÉTODO GENERAL DE CONSTRUCCIÓN DE TEST DE POTENCIA. ESTÁN INCLUIDAS COMO ESCALAS DE CLASIFICACIÓN O DE CATEGORÍAS.		
FUNDAMENTOS DE LA TÉCNICA PRINCIPIOS Y POSTULADOS	- LAS ACTITUDES PUEDEN MEDIRSE A TRAVÉS DE ENUNCIADOS QUE OPEREN COMO REACTIVOS PARA LOS SUJETOS. - LA VARIACIÓN DE LAS RESPUESTAS SERÁ DEBIDA A LAS DIFERENCIAS INDIVIDUALES DE LOS SUJETOS. PUEDEN SITUARSE EN CUALQUIER PUNTO DE LA ESCALA. - LA VARIACIÓN DE LOS SUJETOS EN LA VARIABLE DE ACTITUD NO ES UNIFORME, SINO QUE SUPONE UNA POSICIÓN FAVORABLE O DESFAVORABLE. - ASUME UN NIVEL DE MEDIDA ORDINAL (LOS SUJETOS SON ORDENADOS EN LA ESCALA EN FUNCIÓN DE SU POSICIÓN FAVORABLE/DESFAVORABLE RESPECTO A LA ACTITUD MEDIDA) - LA PUNTUACIÓN OBTENIDA POR LOS SUJETOS EN LA ESCALA ES FUNCIÓN DE LAS PUNTUACIONES OBTENIDAS EN CADA UNO DE LOS ÍTEMS. ES UN MODELO SUMATIVO. ESTO SUPONE ADEMÁS: - LOS ÍTEMS ESTÁN MONOTÓNICAMENTE RELACIONADOS CON EL RASGO (FUNCIÓN MONOTÓNICA Y APROXIMADAMENTE LINEAL RESPECTO A LA ACTITUD MEDIDA) - TODOS LOS ELEMENTOS MIDEN UNA ÚNICA DIMENSIÓN (ESCALA UNIDIMENSIONAL) - LA REDACCIÓN Y PRESENTACIÓN DE LOS ÍTEMS HA DE PERMITIR A LOS SUJETOS EMITIR JUICIOS DE VALOR Y NO JUICIOS DE HECHO. EXPRESAR LO QUE <i>DEBERÍA SER</i> Y NO <i>LO QUE DE HECHO SEA</i>. DEBE SER PRESENTADO DE FORMA QUE CADA SUJETO PUEDA TOMAR PARTIDO ENTRE ALTERNATIVAS OPUESTAS. NORMALMENTE SE RESPONDE EN FUNCIÓN DE 5 CATEGORÍAS (COMPLETAMENTE DE ACUERDO, DE ACUERDO, INDIFERENTE, EN DESACUERDO, COMPLETAMENTE EN DESACUERDO) PERO EL NÚMERO DE OPCIONES DEPENDE DE LO QUE PRETENDA EL INVESTIGADOR, DE LA NATURALEZA DE LA VARIABLE Y DEL TIPO DE ÍTEMS QUE SE ESTÉN UTILIZANDO.	
ASIGNACIÓN DE VALORES NUMÉRICOS A LOS ÍTEMS Y PUNTUACIONES A LOS SUJETOS	SE ASIGNA UN VALOR NUMÉRICO A CADA CATEGORÍA DE RESPUESTA (AL ARBITRIO DEL INVESTIGADOR) - SE DEBE MANTENER UNA COHERENCIA INTERNA: ES NECESARIO QUE SIEMPRE EL VALOR MAS ALTO INDIQUE UNA ACTITUD MÁS POSITIVA (COMPLETAMENTE EN DESACUERDO: 1, -2; EN DESACUERDO: 2, -1; INDIFERENTE: 3, 0; DE ACUERDO: 4, 1; COMPLETAMENTE DE ACUERDO: 5, 2) - LA PUNTUACIÓN DE LOS SUJETOS EN LA ESCALA SERÁ LA SUMA DE LOS VALORES NUMÉRICOS DE LOS ÍTEMS ELEGIDOS.	

4. TÉCNICA DE GUTTMAN: EL ESCALOGRAMA

- MODELO ALTERNATIVO A LA DE THURSTONE Y LIKERT PARA LA **MEDIDA DE ACTITUDES**.
- PUEDE SER UTILIZADO PARA LA CONSTRUCCIÓN DE TEST EN LOS QUE HAY **RESPUESTAS CORRECTAS O INCORRECTAS**.
- MÁS ENCAMINADO A **PROBAR LA UNIDIMENSIONALIDAD DE LA ESCALA** QUE AL PROCESO DE CONSTRUCCIÓN DE LA MISMA: **¿QUÉ GRADO DE DESVIACIÓN RESPECTO A LA ESCALA PERFECTA SE DEBE TOLERAR PARA ACEPTAR QUE SE AJUSTA AL MODELO DE GUTTMAN?**
- SE BASA EN LA IDEA DE QUE **ES POSIBLE ORDENAR LOS ESTÍMULOS DE MANERA QUE SI UN SUJETO RESPONDE CORRECTAMENTE A UN ESTÍMULO CONCRETO, LO HARÁ TAMBIÉN A TODOS LOS QUE ESTÉN SITUADOS POR DEBAJO** DE DICHO ESTÍMULO EN LA ESCALA ESTABLECIDA. LOS ÍTEMS SE ELIGEN SEGÚN SU **GRADO DE EXTREMOSIDAD O DIFICULTAD** (SI LA ESCALA SE CUMPLE EL QUE CONTESTA A UN ÍTEM EXTREMO DEBERÁ HACERLO TAMBIÉN A LOS ANTERIORES)
- CUANDO TENEMOS LA **ESCALA PERFECTA**, DA LUGAR UNA **MATRIZ TRIANGULAR DE 1 Y 0 (ELEMENTOS DICOTÓMICOS: ACIERTO/FALLO)**
- **TANTO LOS SUJETOS COMO LOS ESTÍMULOS** PUEDEN REPRESENTARSE A LO LARGO DE UN **CONTINUO** FORMANDO UNA **ESCALA DE ENTRELAZAMIENTO**: ORDENACIÓN TANTO DE LOS SUJETOS COMO DE LOS ESTÍMULOS RESPECTO A UNA DIMENSIÓN DETERMINADA. COLOCAMOS A LOS **SUJETOS** EN ORDEN DESDE LA PUNTUACIÓN MÁS BAJA A LA MÁS ALTA. Y A SU IZQUIERDA LOS **ÍTEMS** QUE HAN CONTESTADO CORRECTAMENTE; TENEMOS QUE ENTRELAZAR SUJETOS E ÍTEMS PARA QUE TENGA SENTIDO:

EL PATRÓN IDEAL INFERIOR: **1 D 2 C 3 B 4 A**

EVALUACIÓN DEL **ERROR**

- SE LLAMARÁ **ERROR** A LA **DESVIACIÓN DEL PATRÓN DE RR OBSERVADO RESPECTO AL PATRÓN DE RR IDEAL** (NÚMERO DE CAMBIOS QUE HABRÍA QUE HACER PARA AJUSTARNOS AL MODELO PERFECTO, **SIEMPRE ES UN NÚMERO PAR**)
- Nº PATRONES DE R = 2ⁿ**, siendo n nº de ítems de la escala

Sujetos	patrón Rs	Aciertos	patrón ideal	Errores
A	1111	4	1111	0
B	0111	3	1110	2
C	0011	2	1100	4
D	0100	1	1000	2

EN ESTE CASO EL NÚMERO DE PATRONES DE RESPUESTA DISTINTOS ES $2^4 = 16$; PERO DE ESOS POSIBLES 16, SÓLO HAY 4 QUE SE AJUSTAN AL MODELO DE GUTTMAN: (1 1 1 1), (1 1 1 0), (1 1 0 0), (1 0 0 0).

PASOS A SEGUIR PARA LA ELABORACIÓN DE UNA **ESCALA**

- **ORDENAR LA MATRIZ DE RR AJUSTÁNDOLA AL MÁXIMO AL MODELO**:
 1. SUJETOS QUE RESPONDEN CORRECTAMENTE A CADA ÍTEM → ORDENAR LAS FILAS.
 2. PUNTUACIÓN DE CADA SUJETO → INTERCAMBIAR LAS COLUMNAS.
- **RECuento DE ERRORES**: PERMITE VER LA **BONDAD DE AJUSTE DE LOS DATOS AL MODELO**. SE UTILIZA COMO CRITERIO DE BONDAD DE AJUSTE EL **COEFICIENTE DE REPRODUCTIVIDAD (CR)**

$$CR = 1 - \frac{\text{Errores}}{n^{\circ}\text{Respuestas}} = 1 - \frac{\text{Errores}}{N \cdot \text{Ítems}}$$

SI **CR ≥ 0.90** HAY UN **BUEN AJUSTE**

MATRIZ DE RR

	Ítems								
Sujetos	1	2	3	4	5	6	7	8	Pto
A	1	1	1	1	1	1	1	1	8
B	0	0	0	0	0	0	0	0	0
C	1	1	1	0	1	1	0	0	5
D	1	0	0	0	1	0	0	0	2
E	1	1	1	1	1	1	1	1	8
F	1	1	1	0	0	0	0	0	3
G	1	1	1	1	1	0	1	0	6
H	0	0	0	1	0	0	0	0	1
I	1	0	0	0	0	0	0	0	1
J	1	1	1	1	0	0	1	1	6
ACIERTOS	8	6	6	5	5	3	4	3	

MATRIZ DE RR ORDENADA

	Ítems								
Sujetos	1	3	2	5	4	7	6	8	Pto
A	1	1	1	1	1	1	1	1	8
E	1	1	1	1	1	1	1	1	8
G	1	1	1	1	1	1	0	0	6
J	1	1	1	0	1	1	0	1	6
C	1	1	1	1	0	0	1	0	5
F	1	1	1	0	0	0	0	0	3
D	1	0	0	1	0	0	0	0	2
I	1	0	0	0	0	0	0	0	1
H	0	0	0	0	1	0	0	0	1
B	0	0	0	0	0	0	0	0	0
ACIERTOS	8	6	6	5	5	4	3	3	

	Ítems									
Sujetos	1	3	2	4	5	7	8	6	Pto	Errores
A	1	1	1	1	1	1	1	1	8	0
E	1	1	1	1	1	1	1	1	8	0
G	1	1	1	1	1	1	0	0	6	0
J	1	1	1	1	0	1	1	0	6	2
C	1	1	1	0	1	0	0	1	5	2
F	1	1	1	0	0	0	0	0	3	0
D	1	0	0	0	1	0	0	0	2	2
I	1	0	0	0	0	0	0	0	1	0
H	0	0	0	1	0	0	0	0	1	2
B	0	0	0	0	0	0	0	0	0	0
ACIERTOS	8	6	6	5	5	4	3	3		

DIFERENCIAS ENTRE LAS DISTINTAS TÉCNICAS DE ESCALAMIENTO

	THURSTONE	LIKERT	GUTTMAN EL ESCALOGRAMA
MEDIDA DE	ACTITUDES UNIDIMENSIONAL	ACTITUDES UNIDIMENSIONAL	<u>ACTITUDES O COGNITIVAS</u>
TIPO DE ESCALA	ESCALAMIENTO DE <u>ESTÍMULOS PSICOLÓGICOS</u> (ÍTEMS) → DESPUÉS ESCALA A LOS SUJETOS	ESCALAMIENTO DE SUJETOS	ESCALAMIENTO DE <u>RESPUESTAS</u> (ESTÍMULOS Y SUJETOS)
FORMATO DE LOS ÍTEMS	ENUNCIADOS O FRASES Y ALTERNATIVAS.	ENUNCIADOS O FRASES Y ALTERNATIVAS	ENUNCIADOS O FRASES Y ALTERNATIVAS
SELECCIÓN ÍTEMS DEFINITIVOS	VALORES ASIGNADOS EN LA <u>PRUEBA DE JUECES</u>	ES NECESARIO APLICARLA A MUESTRA DE SUJETOS	ES NECESARIO APLICARLA A UNA MUESTRA DE SUJETOS
RESPUESTA	ALTERNATIVAS DICOTÓMICAS SI/NO	VARIAS ALTERNATIVAS. POSICIÓN FAVORABLE/DESF.	
ÍTEMS EN EL CONTÍNUO	EL CONTINUO TIENE UNA DIRECCIÓN CLARA (DE MAYOR A MENOR GRADO DE ACTITUD)	EL CONTINUO PRESENTA UN CAMBIO EN LA DIRECCIÓN DE LA ACTITUD DE FAVORABLE A DESF.	
	ES NECESARIO INCLUIR ÍTEMS QUE CUBRAN TODO EL CONTINUO DE LA ACTITUD	LOS ENUNCIADOS DE LOS ÍTEMS SE SITUAN SOLO EN LOS <u>DOS EXTREMOS DEL CONTINUO</u> .	ES NECESARIO INCLUIR ÍTEMS QUE CUBRAN TODO EL CONTINUO DE LA ACTITUD
	NO SON ACUMULATIVAS	NO SON ACUMULATIVAS	SON <u>ACUMULATIVAS</u> . ORDEN ÍTEMS POR GRADO DE EXTREMOSIDAD O DIFICULTAD.
NIVEL DE MEDIDA	<u>ESCALA DE INTERVALO</u>	ESCALA ORDINAL	ESCALA ORDINAL
PUNTUACIÓN EN LA ESCALA	<u>MEDIA</u> DE LOS VALORES ESCALARES DE LOS ÍTEMS SELECCIONADOS.	SON <u>SUMATIVAS</u> : SUMA DE LOS VALORES NUMÉRICOS ÍTEMS.	--
VENTAJAS	INTERPRETACIÓN DIRECTA DE LA ACTITUD DE UN SUJETO SIN REFERENCIA AL GRUPO.	SON LAS MAS UTILIZADAS, SON RÁPIDAS Y SE ADAPTAN A LA MAYORÍA DE LAS ACTITUDES	ORIENTADA A PROBAR LA UNIDAMENSIONALIDAD DE UNA ESCALA O AJUSTE AL MODELO
DESVENTAJAS	EXCESIVA DEPENDENCIA DEL CRITERIO DE LOS JUECES. SE USAN POCO.	FACILIDAD RESPUESTAS FALSAS, INTERVALOS ENTRE PUNTOS DE LA ESCALA NO IGUALES PARA ACTITUD SUJETOS	

RESUMEN MÉTODOS DE ESCALAMIENTO

MÉTODOS CENTRADOS EN LOS SUJETOS	TEORÍA DE LOS TEST LIKERT OSGOOD	
MÉTODOS CENTRADOS EN LOS ESTÍMULOS	ESCALAMIENTO PSICOFÍSICO ESCALAMIENTO PSICOLÓGICO	FECHNER STEVENS THURSTONE
MÉTODOS CENTRADOS EN LAS RESPUESTAS	ESCALOGRAMA DE GUTTMAN MÉTODO DE ESCALA DE COOMBS	

TEMA 4. LA FIABILIDAD DE LAS PUNTUACIONES

EVALUACIÓN DE LA CALIDAD MÉTRICA DE LA PRUEBA PILOTO Y CONSTRUCCIÓN DEL INSTRUMENTO DEFINITIVO: FIABILIDAD Y VALIDEZ

CONSTRUCCIÓN DE TEST COMO INSTRUMENTO CIENTÍFICO DE MEDICIÓN. ETAPAS	1. INICIAL: CONSTRUCCIÓN (TEMAS 2 Y 3) 2. SEGUNDA: EVALUACIÓN (TEMAS 8 Y 4, 5, 6 Y 7) 3. TERCERA: APLICACIÓN Y EVALUACIÓN A LOS SUJETOS (TEMA 9) 4. FINAL: MANUAL DEL TEST (TEMA 9) HASTA AHORA HEMOS TRABAJADO EN LA FASE INICIAL: LOS PRINCIPIOS BÁSICOS PARA LA CONSTRUCCIÓN DE TESTS Y LAS PRINCIPALES TÉCNICAS DE CONSTRUCCIÓN DE ESCALAS DE ACTITUDES. AHORA COMENZAMOS CON LA EVALUACIÓN DE LA CALIDAD MÉTRICA DEL TEST (FIABILIDAD Y VALIDEZ): EVALUACIÓN DE CADA UNO DE LOS ÍTEMS (TEMA 8) + EVALUACIÓN DEL CONJUNTO DEL TEST (TEMAS 4, 5, 6 Y 7)
CONCEPTOS FUNDAMENTALES FIABILIDAD: PRECISION Y ESTABILIDAD DE LAS MEDIDAS. CADA VEZ QUE MIDO CON EL METRO LA MISMA MESA, MIDE LO MISMO. VALIDEZ: PERTINENCIA DE LAS INFERENCIAS REALIZADAS A PARTIR DE LAS PUNTUACIONES OBTENIDAS. QUE EFECTIVAMENTE EL METRO MIDE LA LONGITUD Y NO EL PESO, POR EJEMPLO. PODEMOS TENER UN INSTRUMENTO MUY FIABLE (METRO) PERO QUE NO VALGA PARA LO QUE PRETENDEMOS MEDIR (PESO)	
*EL PROBLEMA DEL ERROR DE MEDIDA UNO DE LOS REQUISITOS FUNDAMENTALES EN CUALQUIER TEORÍA DE LA MEDICIÓN ES LA FIABILIDAD Y PRECISIÓN DE LOS INSTRUMENTOS DE MEDIDA. SI APLICÁRAMOS n VECES UN TEST A UN SUJETO, ES CASI SEGURO QUE LAS PUNTUACIONES OBTENIDAS VARIARÁN. ESTOS ERRORES NO PUEDEN SER DEBIDOS AL INSTRUMENTO DE MEDIDA (TEST) SINO QUE SON ERRORES DE CARÁCTER ALEATORIO E IMPREDECIBLES DEBIDO A CAMBIOS QUE OPERAN EN EL PROPIO SUJETO (MOTIVACIÓN, AZAR, CONDICIONES FÍSICAS...) SON DE LOS QUE SE OCUPA LA FIABILIDAD. ¿CÓMO PODEMOS SABER CUÁL ES EL VALOR REAL DEL SUJETO EN LA CARACTERÍSTICA QUE ESTAMOS ESTUDIANDO? → ESTIMACIONES DE LA CUANTÍA DE ERROR A TRAVÉS DEL MODELO LINEAL DE SPEARMAN.	

1. MODELO LINEAL DE SPEARMAN (TEORÍA CLÁSICA DEL TEST)

LA PUNTUACIÓN EMPÍRICA OBTENIDA POR UN SUJETO EN UN TEST (X) ES UNA COMBINACIÓN LINEAL DE SU PUNTUACIÓN VERDADERA (V) Y UN ERROR DE MEDIDA (E). *ERROR DE MEDIDA → DIFERENCIA ENTRE LA PUNTUACIÓN EMPÍRICA OBTENIDA (OBSERVADA) POR UN SUJETO EN UN TEST Y SU PUNTUACIÓN VERDADERA.	$X = V + E$	SUPUESTOS 1. $V = E(X)$ 2. $r_{ve} = 0$ 3. $r_{e1e2} = 0$	DEDUCCIONES: 1. $E = X - V$ 2. $E(e) = 0$ 3. $\bar{X} = \bar{V}$ 4. $Cov(V, E) = 0$ 5. $S_x^2 = S_v^2 + S_e^2$ 6. $Cov(X, V) = S_v^2$ 7. $r_{xe} = \frac{S_e}{S_x}$ 8. $Cov(X_1, X_2) = Cov(V_1, V_2)$
--	-------------	---	--

2. TEST PARALELOS. CONDICIONES DE PARALELISMO

AQUELLOS QUE MIDEN LA MISMA VARIABLE PERO ESTÁN FORMADOS POR DISTINTOS ÍTEMS. (WAIS – BINET) SI A UNA MISMA MUESTRA DE SUJETOS SE LE APLICAN DOS TEST, X Y X' , CONSIDERAMOS QUE SON PARALELOS SI, ADÉMÁS DE CUMPLIRSE LOS SUPUESTOS ANTERIORES, SE CUMPLEN LAS DOS CONDICIONES SIGUIENTES: *LAS PUNTUACIONES VERDADERAS Y LA VARIANZA DE LOS ERRORES SON LAS MISMAS RESPECTIVA/	$V = V'$ $S_e^2 = S_{e'}$	DEDUCCIONES: 1. $\bar{X} = \bar{X}'$ 2. $S_x^2 = S_{x'}$ 3. $r_{xx'} = r_{xv}^2 = \frac{S_v^2}{S_x^2}$ 4. $r_{x1x2} = r_{x1x3} = r_{x2x3} = \dots = r_{xjxk}$
---	---------------------------	---

3. INTERPRETACIÓN TEÓRICA DEL COEFICIENTE DE FIABILIDAD

COEFICIENTE DE FIABILIDAD $r_{xx'}$ CORRELACION ENTRE LAS PUNTUACIONES EMPÍRICAS OBTENIDAS POR UNA MUESTRA DE SUJETOS EN DOS FORMAS PARALELAS DEL TEST. *PROPORCIÓN DE LA VARIANZA VERDADERA QUE HAY EN LA VARIANZA EMPÍRICA.	$r_{xx'} = \frac{S_v^2}{S_x^2}$	$r_{xx'} = 1 \rightarrow$ FIABILIDAD PERFECTA $r_{xx'} = 0 \rightarrow S_e^2 = S_x^2$ VARIANZA DE ERRORES IGUAL A VARIANZA EMPÍRICA
	$r_{xx'} = 1 - \frac{S_e^2}{S_x^2} = 1 - r_{xe}^2$	$r_{xe} = \frac{S_e}{S_x} \Rightarrow r_{xe} = \sqrt{1 - r_{xx'}}$
ÍNDICE DE FIABILIDAD DE UN TEST r_{xv}	$r_{xv} = \sqrt{r_{xx'}} \rightarrow$	$r_{xv} \geq r_{xx'}$ ÍNDICE DE FIABILIDAD SIEMPRE $\geq$ QUE COEF. FIABILIDAD

4. TIPOS DE ERRORES DE MEDIDA

	ERROR PARTICULAR. DIF. ENTRE 2 PUNT.	MEDIDA GRUPAL DEL ERROR	
	ERROR	ERROR TÍPICO EN PUNT. DIRECTAS Y DIFERENCIALES	ERROR TÍPICO EN PUNTUACIONES TÍPICAS
DE MEDIDA DIFERENCIA ENTRE LA <u>PUNTUACIÓN EMPÍRICA</u> DE UN SUJETO Y LA <u>VERDADERA</u> .	$E = X - V$	$S_e = S_x \sqrt{1 - r_{xx'}}$	$S_{Z_e} = \sqrt{1 - r_{xx'}}$
DE ESTIMACIÓN DE LA PUNTUACIÓN VERDADERA DIFERENCIA ENTRE <u>PUNTUACIÓN VERDADERA (V)</u> Y <u>PUNTUACIÓN VERDADERA PRONOSTICADA (V')</u> MEDIANTE EL MODELO DE REGRESIÓN.	$E = V - V'$	$S_{vx} = S_x \sqrt{1 - r_{xx'} \sqrt{r_{xx'}}$ $S_{vx} = S_e \sqrt{r_{xx'}}$	$S_{Z_{vx}} = \sqrt{1 - r_{xx'} \sqrt{r_{xx'}}$ $S_{Z_{vx}} = S_{Z_e} \sqrt{r_{xx'}}$
DE SUSTITUCIÓN DIFERENCIA ENTRE LAS <u>PUNTUACIONES OBTENIDAS</u> POR UN SUJETO EN UN <u>TEST (X₁)</u> Y LAS OBTENIDAS EN LA FORMA PARALELA DEL <u>TEST (X₂)</u> . ERROR COMETIDO SI SUSTITUIMOS LAS PUNTUACIONES DE UN TEST X ₁ POR UN TEST X ₂ .	$e = X_1 - X_2$	$S_{x_1-x_2} = S_x \sqrt{1 - r_{xx'} \sqrt{2}}$ $S_{x_1-x_2} = S_e \sqrt{2}$	$S_{x_1-x_2} = \sqrt{1 - r_{xx'} \sqrt{2}}$ $S_{x_1-x_2} = S_{Z_e} \sqrt{2}$
DE PREDICCIÓN DIFERENCIA ENTRE LAS <u>PUNTUACIONES OBTENIDAS</u> POR UN SUJETO EN UN <u>TEST (X₁)</u> Y LAS <u>PUNTUACIONES PRONOSTICADAS (X'₁)</u> A PARTIR DE LA FORMA PARALELA DEL <u>TEST (X₂)</u> . r_{12} COEF. FIABILIDAD: CORRELACIÓN ENTRE LAS FORMAS PARALELAS.	$e = X_1 - X'_1$ SIENDO $X'_1 = r_{12} \frac{S_{x_1}}{S_{x_2}} (X_2 - \bar{X}_2) + \bar{X}_1$	$S_{ep} = S_x \sqrt{1 - r_{xx'} \sqrt{1 + r_{xx'}}$ $S_{ep} = S_e \sqrt{1 + r_{xx'}}$	$S_{Z_{ep}} = \sqrt{1 - r_{xx'} \sqrt{1 + r_{xx'}}$ $S_{Z_{ep}} = S_{Z_e} \sqrt{1 + r_{xx'}}$

5. FACTORES QUE AFECTAN A LA FIABILIDAD DEL TEST

LONGITUD	A MEDIDA QUE <u>AUMENTA LA LONGITUD</u> DEL TEST SE <u>INCREMENTA EL COEFICIENTE DE FIABILIDAD</u> . NO SE DA UN CRECIMIENTO PROPORCIONAL, ES UNA FUNCIÓN MONÓTONA CRECIENTE: <u>¿HASTA QUE PUNTO ES RAZONABLE AUMENTARLO?</u> n ES EL NÚMERO DE VECES QUE SE HA ALARGADO (O DISMINUIDO EL TEST) $n = \frac{n^{\circ} \text{ elementos finales}}{n^{\circ} \text{ elementos iniciales}}$	<u>ECUACIÓN DE SPEARMAN-BROWN</u>	$R_{xx'} = \frac{nr_{xx'}}{1 + (n - 1)r_{xx'}}$
		SI LA LONGITUD ES EL DOBLE:	$R_{xx'} = \frac{2r_{xx'}}{1 + r_{xx'}}$
			CÁLCULO DEL Nº VECES QUE AUMENTA O DISMINUYE TEST
VARIABILIDAD	<u>CUANTO MÁS HOMOGÉNEO SEA EL GRUPO, MENOR SERÁ EL COEFICIENTE DE FIABILIDAD</u> . UN TEST PRESENTA TANTOS COEFICIENTES DE FIABILIDAD COMO MUESTRAS DISTINTAS EN LAS QUE SE CALCULE. <u>SI SUPONEMOS DOS GRUPOS DE SUJETOS 1 Y 2:</u>	$r_{22} = 1 - \frac{S_1^2}{S_2^2}(1 - r_{11})$	
CARACTERÍSTIMAS DE LOS ÍTEMAS (SE ABORDAN EN EL TEMA 8)			

6. CÁLCULO DEL COEFICIENTE DE FIABILIDAD

FORMAS PARALELAS: <u>COEFICIENTE DE EQUIVALENCIA</u> $r_{xx'} = r_{x_1x_2}$	<u>CORRELACIÓN ENTRE LAS DOS FORMAS PARALELAS DEL TEST</u> . VENTAJAS: SI AMBAS FORMAS SON APLICADAS EN EL MISMO MOMENTO SE TIENE UN MAYOR CONTROL DE LAS CONDICIONES EN QUE LOS SUJETOS REALIZAN LAS PRUEBAS. INCONVENIENTE: LA PROPIA CONSTRUCCIÓN DE DOS FORMAS PARALELAS.	COEFICIENTE DE CORRELACIÓN DE PEARSON: $r_{x_1x_2} = \frac{n \sum X_1 X_2 - \sum X_1 \sum X_2}{\sqrt{n \sum X_1^2 - (\sum X_1)^2} \sqrt{n \sum X_2^2 - (\sum X_2)^2}}$
TEST - RETEST: <u>COEFICIENTE DE ESTABILIDAD</u> $r_{x_1x'_1} = r_{x_1x_2}$	<u>CORRELACIÓN ENTRE LA PRIMERA Y SEGUNDA APLICACIÓN DE UN MISMO TEST A LA MISMA MUESTRA</u> . VENTAJAS: NO REQUIERE LA CONSTRUCCIÓN DE FORMAS PARALELAS. INCONVENIENTE: MEMORIZACIÓN DE ALGUNOS ÍTEMAS, EXCESIVO INTERVALO DE TIEMPO ENTRE UNA APLICACIÓN Y OTRA EN RASGOS QUE VARÍAN Y CAMBIO EN EL GRADO DE COOPERACIÓN DEL SUJETO.	

7. LA FIABILIDAD COMO CONSISTENCIA INTERNA

CUANDO SÓLO PODEMOS APLICAR EL TEST UNA VEZ PODEMOS UTILIZAR LOS SIGUIENTES MÉTODOS:

MÉTODOS BASADOS EN LA DIVISIÓN DEL TEST EN DOS MITADES

* LOS TEST DEBEN SER SIMILARES EN CONTENIDO Y DIFICULTAD.
DIVISIÓN:
- PRIMEROS $n/2$, ÚLTIMOS $n/2$
- PARES/IMPARES
- ORDEN POR GRADO DE DIFICULTAD+PARES/IMPARES
- AZAR

1. COMPROBAMOS CONDICIONES DE PARALELISMO
2. CORRELACIÓN ENTRE LAS DOS MITADES EN LAS QUE DIVIDIMOS EL TEST (ÍTEMS PARES E IMPARES):

$$r_{pi} = \frac{n \sum X_p X_i - \sum X_p \sum X_i}{\sqrt{n \sum X_p^2 - (\sum X_p)^2} \sqrt{n \sum X_i^2 - (\sum X_i)^2}}$$

3. APLICAMOS FÓRMULAS DE CORRECCIÓN.

VENTAJAS: APLICACIÓN ÚNICA DEL TEST (EVITAMOS APRENDIZAJE, CAMBIO POR EL PASO DEL TIEMPO, MEMORIZACIÓN, ETC) Y AHORRO DE TIEMPO Y ESFUERZO AL EVITAR LA CONSTRUCCIÓN DE FORMA PARALELA DEL TEST O REEVALUACIÓN DE LOS SUJETOS.

PARA TEST PARALELOS ($V = V'$, $S_e^2 = S_{e'}^2$)

FÓRMULA DE SPEARMAN-BROWN:

$$R_{xx'} = \frac{2r_{pi}}{1 + r_{pi}}$$

PARA TEST τ -EQUIVALENTES ($V = V'$, $S_e^2 \neq S_{e'}^2$)

FÓRMULA DE RULON:

$$r_{xx'} = 1 - \frac{S_d^2}{S_x^2} = 1 - \frac{S_{p-i}^2}{S_x^2}$$

$$S_d^2 = \frac{\sum d^2}{N} - \left(\frac{\sum d}{N} \right)^2 \text{ O BIEN:}$$

$$S_d^2 = S_{p-i}^2 = S_p^2 + S_i^2 - 2r_{pi} S_p S_i$$

o FÓRMULA DE GUTTMAN-FLANAGAN:

$$r_{xx'} = 2 \left[1 - \frac{S_p^2 + S_i^2}{S_x^2} \right] = \frac{4r_{pi} S_p S_i}{S_x^2}$$

MÉTODOS BASADOS EN LA COVARIACIÓN ENTRE LOS ÍTEMS

COEFICIENTE α DE CRONBACH

- EXPRESA LA FIABILIDAD DEL TEST EN FUNCIÓN DEL NÚMERO DE ÍTEMS Y DE LA PROPORCIÓN DE LA VARIANZA TOTAL DEL TEST DEBIDA A LA COVARIACIÓN ENTRE LOS ÍTEMS.
- CUANTO MAS COVARÍEN LOS ÍTEMS ENTRE SI, MAYOR SERÁ LA FIABILIDAD DEL TEST.
- EL ESTIMADOR INESGADO DE α SERÁ $\bar{\alpha}$, (SIENDO $\hat{\alpha}$, α DE CRONBACH, N Nº SUJETOS). CUANDO $N \rightarrow \infty \Rightarrow \bar{\alpha} \sim \hat{\alpha}$, EN LA PRÁCTICA CUANDO $N > 100$.
- EL COEFICIENTE α (TAMBIÉN δ_3) PUEDE SER CONSIDERADO COMO UNA ESTIMACIÓN DEL LÍMITE INFERIOR DEL COEFICIENTE DE FIABILIDAD DE UN TEST, SIENDO SU VALOR MENOR O IGUAL QUE EL COEFICIENTE DE CORRELACIÓN.

$$\alpha \leq r_{xx'}$$

$$\alpha = r_{xx'} \text{ SI TEST PARALELOS}$$

CASOS PARTICULARES DEL COEFICIENTE α PARA ÍTEMS DICOTÓMICOS

LAS KUDER-RICHARDSON:

EN CASO DE APLICAR KR_{21} CON ÍTEMS DE DISTINTA DIFICULTAD, OBTENEMOS UN VALOR INFERIOR A KR_{20} .

1. CÁLCULO DEL COEFICIENTE:

$$\alpha = \frac{n}{n-1} \left(1 - \frac{\sum S_j^2}{S_x^2} \right)$$

2. INFERENCIAS SOBRE α : PARA IR MÁS ALLÁ ESTUDIAMOS SI α ES ESTADÍSTICAMENTE SIGNIFICATIVO O NO EN DIFERENTES CASOS. BUSCAMOS INTERVALO CONFIDENCIAL PARA α EN LA POBLACIÓN.

1. UNA MUESTRA
2. MUESTRAS DEPENDIENTES
3. MUESTRAS INDEPENDIENTES

DISTINTA DIFICULTAD:

$$KR_{20} = \frac{n}{n-1} \left(1 - \frac{\sum p_j q_j}{S_x^2} \right)$$

MISMA DIFICULTAD:

$$KR_{21} = \frac{n}{n-1} \left(1 - \frac{npq}{S_x^2} \right)$$

MÉTODOS BASADOS EN EL ANÁLISIS FACTORIAL DE LOS ÍTEMS

APROXIMACIONES A α

COEFICIENTE THETA DE CARMINES Y ZELLER ϑ

COEFICIENTE OMEGA DE HEISE Y BOHRNSTEDT Ω

AMBOS SON INDICADORES DE LA CONSISTENCIA INTERNA DEL TEST Y APROXIMACIONES AL COEFICIENTE α .

ϑ TAMBIÉN ES UN INDICADOR DE LA UNIDIMENSIONALIDAD DE LOS ÍTEMS.

EN GENERAL, $\alpha \leq \vartheta \leq \Omega$, PERO SI ÍTEMS $\parallel \rightarrow \alpha = \vartheta = \Omega$

COEFICIENTE BETA DE RAJÚ

CUANDO TENEMOS UN TEST COMPUESTO POR VARIOS SUBTEST CON DISTINTO NÚMERO DE ÍTEMS, Y DESCONOCEMOS LAS PUNTUACIONES EN LOS DISTINTOS SUBTEST.

SE APLICA ESTE COEFICIENTE CUANDO SE DESCONOCEN LAS PUNTUACIONES DE LOS SUJETOS EN LOS ÍTEMS DE LOS DISTINTOS SUBTEST, PORQUE α NOS DARÍA UN VALOR INFRAESTIMADO.

EN EL CASO DE CONOCERLAS, ES MEJOR EMPLEAR EL COEF ALFA.

$$\theta = 1 - \frac{\sum_{j=1}^n S_j^2 - \sum_{j=1}^n S_j^2 h_j^2}{\sum_{j=1}^n \sum_{h=1}^n \text{Cov}(X_j, X_h)} \quad \text{o} \quad \theta = 1 - \frac{n - \sum h_j^2}{n+2 \sum r_{jh}}$$

$$\Omega = \frac{n}{n-1} \left(1 - \frac{1}{\lambda_1} \right)$$

$$\beta = \frac{S_x^2 - \sum S_j^2}{S_x^2 \left(1 - \sum \left(\frac{n_j}{N} \right)^2 \right)}$$

INFERENCIAS PARA α

1. Una muestra

$$H_0: \alpha = \alpha_0$$

El coeficiente α NO es estadísticamente significativo

$$H_1: \alpha \neq \alpha_0$$

El coeficiente α es estadísticamente significativo

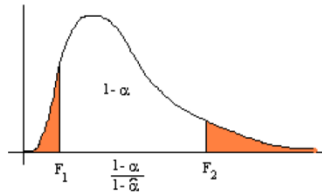
$$F = \frac{1 - \alpha}{1 - \tilde{\alpha}}$$

Se distribuye según F con (N-1) y (n-1) (N-1) grados de libertad

✓ Intervalo de Confianza

$$l_i = 1 - \left[F_{\alpha/2}(1 - \hat{\alpha}) \right]$$

$$l_s = 1 - \left[F_{1-\alpha/2}(1 - \hat{\alpha}) \right]$$



2. Dos muestras dependientes

Se quiere comprobar si dos coeficientes α son estadísticamente significativos o no entre si en la misma muestra

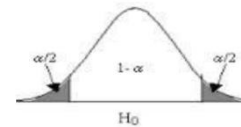
$$H_0: \alpha_1 = \alpha_2$$

$$H_1: \alpha_1 \neq \alpha_2$$

Estadístico de Contraste

$$T = \frac{(\hat{\alpha}_1 - \hat{\alpha}_2)\sqrt{N-2}}{\sqrt{4(1-\hat{\alpha}_1)(1-\hat{\alpha}_2)(1-r_{12}^2)}}$$

Distribución T de Student con $N-2$ grados de libertad



3. Dos muestras independientes

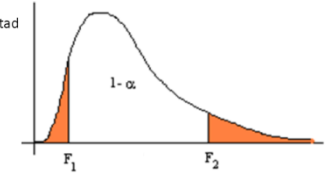
Se quiere comprobar si la fiabilidad de un test (α) es estadísticamente significativo o no entre si en distintas muestras

$$H_0: \alpha_1 = \alpha_2$$

$$H_1: \alpha_1 \neq \alpha_2$$

Estadístico de Contraste $W = \frac{1 - \alpha_1}{1 - \alpha_2}$

Distribución F con $N_1 - 1$ y $N_2 - 1$ grados de libertad



9. CÁLCULO DE LA PUNTUACIÓN VERDADERA

	BASADA EN LA DISTRIBUCIÓN NORMAL DE LOS ERRORES	BASADA EN EL MODELO DE REGRESIÓN	MEDIANTE LA DESIGUALDAD DE CHEBYCHEV
Valor crítico	$Z_{\alpha/2}$	$Z_{\alpha/2}$	$K = \sqrt{\frac{1}{\alpha}}$
Error típico	$S_e = S_x \sqrt{1 - r_{xx}}$	$S_{vx} = S_x \sqrt{1 - r_{xx}} \sqrt{r_{xx}}$ $= S_e \sqrt{r_{xx}}$	$S_e = S_x \sqrt{1 - r_{xx}}$
Error máximo	$E_{max} = Z_{\alpha/2} S_e$	$E_{max} = Z_{\alpha/2} S_{vx}$	$E_{max} = K S_e$
DIRECTAS X	$X \pm E_{max}$	$V' = r_{xx}(X - \bar{X}) + \bar{X}$ $V' \pm E_{max}$	$X \pm E_{max}$
DIFERENCIALES	$x \pm E_{max}$	$v' = r_{xx} x$ $v' \pm E_{max}$	

10. CÁLCULO FIABILIDAD BATERÍA DE TEST

SE CALCULA EN FUNCIÓN DE LOS COEFICIENTES DE FIABILIDAD, VARIANZAS Y COVARIANZAS DE LOS SUBTEST QUE LA VAN A CONFORMAR:

$$r_{tt} = 1 - \frac{\sum S_j^2 - \sum S_j^2 r_{jj}}{S_t^2}$$

SIENDO:

S_j^2 VARIANZA DEL SUBTEST j

r_{jj} COEFICIENTE DE FIABILIDAD DEL SUBTEST j

S_t^2 VARIANZA DE LA BATERÍA TOTAL

TEMA 6. VALIDEZ DE LAS INFERENCIAS I

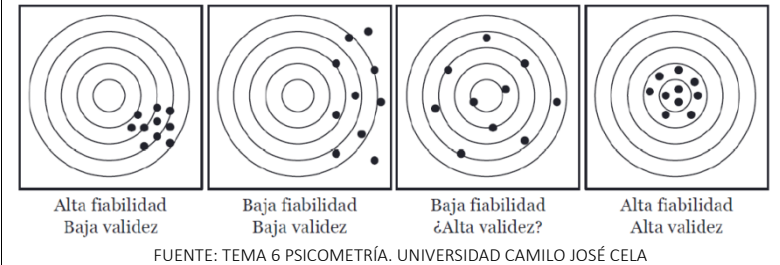
¿QUÉ ES LA VALIDEZ?

FIABILIDAD: PRECISION Y ESTABILIDAD DE LAS MEDIDAS. GRADO EN QUE LAS PUNTUACIONES DE UN TEST REFLEJAN UN NIVEL DE RASGO O COMPETENCIA, ES DECIR, ESTÁN LIBRES DE LOS ERRORES ALEATORIOS PRESENTES EN CUALQUIER PROCESO DE MEDICIÓN. CADA VEZ QUE MIDO CON EL METRO LA MISMA MESA, MIDE LO MISMO.

VALIDEZ: PERTINENCIA DE LAS INFERENCIAS REALIZADAS A PARTIR DE LAS PUNTUACIONES OBTENIDAS. GRADO DE RELACIÓN ENTRE LA EVIDENCIA EMPÍRICA OBTENIDA Y EL CONCEPTO TEÓRICO DEL CONSTRUCTO QUE SE INTENTA MEDIR. QUE EFECTIVAMENTE EL METRO MIDE LA LONGITUD Y NO EL PESO, POR EJEMPLO. PODEMOS TENER UN INSTRUMENTO MUY FIABLE (METRO) PERO QUE NO VALGA PARA LO QUE PRETENDEMOS MEDIR (PESO)

LA **FIABILIDAD** DE UN INSTRUMENTO DE MEDIDA ES UNA **CONDICIÓN NECESARIA PERO NO SUFICIENTE**. DEBEMOS PLANTEARNOS HASTA QUÉ PUNTO EL INSTRUMENTO DE MEDIDA SIRVE PARA LA FINALIDAD PERSEGUIDA CON SU APLICACIÓN, PUESTO QUE **DEBE PROPORCIONAR MEDIDAS FIABLES A PARTIR DE LAS CUALES SE PUEDAN HACER INFERENCIAS VÁLIDAS**.

FIABILIDAD Y VALIDEZ NO SON **PROPIEDADES INTRÍNSECAS DEL TEST SINO DE LOS OBJETIVOS DE SU USO Y FINES**. LA VALIDEZ NO SE REFIERE AL TEST EN SÍ MISMO SINO AL BUEN USO DE LOS RESULTADOS. **NO SE VALIDA EL INSTRUMENTO SINO, COMO SEÑALA CORBACHO (1971), LA INTERPRETACIÓN DE LOS DATOS** OBTENIDOS POR MEDIO DE UN PROCEDIMIENTO ESPECÍFICO.



TIPOS DE VALIDEZ

LOS TIPOS DE VALIDEZ VAN UNIDOS A OBJETIVOS CONCRETOS EN EL USO DE LOS TEST (APA 1966)



VALIDEZ DE CRITERIO O CORRELACIONAL: CAPACIDAD DEL TEST PARA PREDECIR/ESTIMAR UN CRITERIO EXTERNO. LA FORMA DE OPERATIVIZAR LA RELACIÓN ENTRE TEST Y CRITERIO ES MEDIANTE UN COEFICIENTE DE CORRELACIÓN.

- **PREDICTIVA**: CRITERIO POSTERIOR AL TEST (SELECCIÓN Y CLASIFICACIÓN DE PERSONAL)
- **CONCURRENTE**: LA RECOGIDA DE INFORMACIÓN, TANTO DEL TEST COMO DEL CRITERIO, SE HACE SIMULTÁNEAMENTE (DIAGNÓSTICO)
- **RETROSPECTIVA**: MEDIDA DEL CRITERIO CON ANTERIORIDAD AL TEST (ANÁLISIS/PSICOMETRÍA)

CONTENIDO: RELEVANCIA Y REPRESENTATIVIDAD DE LOS ÍTEMS. CRITERIOS INTERNOS AL PROPIO TEST (EL TEST EN SÍ MISMO CONSTITUYE SU PROPIO CRITERIO, TEST DE CONOCIMIENTOS)

CONSTRUCTO: MEDIDAS DEL MISMO CONSTRUCTO OBTENIDAS CON OTRAS PRUEBAS. ¿QUÉ ES LO QUE MIDE REALMENTE EL TEST? ¿MIDE SÓLO ESO?: OBTENER LA GARANTÍA DE QUE LAS CONDUCTAS OBSERVABLES QUE SE HAN ELEGIDO COMO INDICADORES DEL CONSTRUCTO, LO SON REALMENTE.

DESDE EL PUNTO DE VISTA CIENTÍFICO LA VALIDEZ QUE HAY QUE CONSIDERAR ES LA DE CONSTRUCTO, LAS OTRAS QUEDARÍAN INCLUIDAS EN ÉSTA Y SE CONSIDERAN **ESTRATEGIAS DE VALIDACIÓN** PARA COMPRENDER MEJOR LO QUE MIDE UN TEST. EL **PROCESO DE VALIDACIÓN** SERÁ UN PROCESO CONTINUO QUE PERMITIRÁ OBTENER DISTINTOS TIPOS DE EVIDENCIA EMPÍRICA QUE IMPLICAN LOS 3 TIPOS DE VALIDEZ, EN FUNCIÓN DE LA INTERPRETACIÓN QUE SE VAYA A HACER DE LAS PUNTUACIONES OBTENIDAS Y EL OBJETIVO QUE SE PRETENDA ALCANZAR, ES DECIR, DE LAS INFERENCIAS QUE SE VAYAN A HACER.

ACTUALMENTE SE CONSIDERA LA **VALIDEZ** COMO UN CONCEPTO UNITARIO QUE HACE REFERENCIA AL **GRADO EN QUE EL TEST MIDE AQUELLO QUE PRETENDE MEDIR** O **GRADO EN QUE LA EVIDENCIA EMPÍRICA OBTENIDA Y LOS CONOCIMIENTOS APORTADOS POR LAS TEORÍAS APOYAN LAS INFERENCIAS QUE SE HAGAN A PARTIR DE LAS PUNTUACIONES OBTENIDAS EN EL TEST CUANDO ÉSTE SE UTILIZA PARA UN OBJETIVO CONCRETO**.

1. VALIDEZ DE CONTENIDO

TIENE QUE VER CON LOS ÍTEMS Y CON LAS INSTRUCCIONES DE ADMINISTRACIÓN, CORRECCIÓN Y PUNTUACIÓN.

EL OBJETIVO ES ANALIZAR E INDICAR SI LOS **ÍTEMS** O ELEMENTOS DEL TEST SON UNA **MUESTRA RELEVANTE Y REPRESENTATIVA DEL CONSTRUCTO** SOBRE EL QUE SE VA HACER INFERENCIAS.

LA FORMA TÍPICA PARA LLEVAR A CABO UN ESTUDIO DE VALIDEZ DE CONTENIDO ES UTILIZAR UN **GRUPO DE EXPERTOS** (JUICIOS SUBJETIVOS) QUE ANALIZARÁN DOS ASPECTOS:

- **ESPECIFICACIÓN DEL DOMINIO DE CONDUCTAS** → ELIMINAR ASPECTOS IRRELEVANTES DEL DOMINIO DE INTERÉS. ES UN **ANÁLISIS CUALITATIVO Y CUANTITATIVO**. (1) **MATRIZ DE DOBLE ENTRADA**: ÁREAS DE CONTENIDO Y PROCESOS (TAXONOMÍA DE BLOOM) QUE SE VAN A EVALUAR Y LA IMPORTANCIA RELATIVA DE CADA UNO DE ELLOS. NOS DA PORCENTAJE DE ÍTEMS DEL TEST. (2) **GRADO DE ACUERDO ENTRE EXPERTOS**: ESCALA DE 5 PUNTOS DE GRADO DE AJUSTE DE CADA ÍTEM CON SU CORRESPONDIENTE ESPECIFICACIÓN DE DOMINIO. DESPUÉS HALLAMOS LA MEDIA O MEDIANA DE LOS VALORES ASIGNADOS. SE SELECCIONAN LOS ÍTEMS CON ALTO GRADO DE AJUSTE.
- **REPRESENTATIVIDAD DEL DOMINIO** → QUE EL TESTS INCLUYA TODOS LOS ELEMENTOS/CONDUCTAS IMPORTANTES QUE DEFINEN EL DOMINIO. GRADO EN EL QUE SE HAN CUBIERTO LAS ESPECIFICACIONES DE DOMINIO, TANTO EN CUANTO A CONTENIDOS COMO A OBJETIVOS.

2. VALIDEZ DE CONSTRUCTO

ES LA VALIDEZ QUE REALMENTE DA **SIGNIFICADO** A LAS PUNTUACIONES DE LOS TEST, PUES **PERMITE OBTENER EVIDENCIA DE QUE LAS CONDUCTAS OBSERVABLES** QUE SE HAN ELEGIDO COMO **INDICADORES DEL CONSTRUCTO** (VARIABLE LATENTE INOBSERVABLE) REALMENTE LO SON.

SE CENTRAN EN EL **ANÁLISIS DE LA ESTRUCTURA DEL TEST**, ES DECIR, DE LAS **INTERRELACIONES** ENTRE:

- LAS PUNTUACIONES EN LOS DISTINTOS ÍTEMS (**ESTRUCTURA INTERNA DEL TEST**)
- LAS RELACIONES ENTRE EL RESULTADO DEL TEST Y OTRAS MEDIDAS DEL MISMO CONSTRUCTO (**ESTRUCTURA EXTERNA DEL TEST**)

PARA ELLO DEBEMOS:



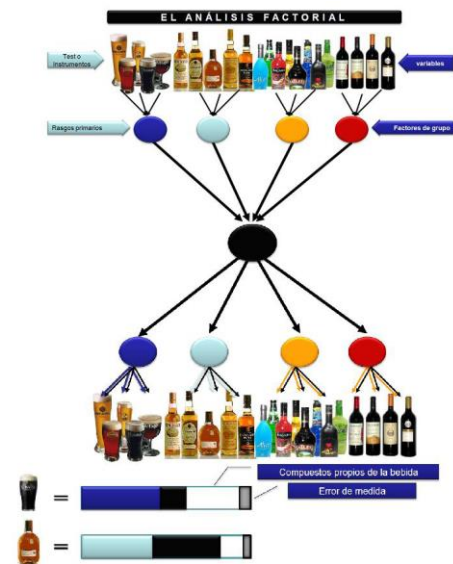
1. **DEFINIR EL CONSTRUCTO** Y SU HIPOTÉTICA **RELACIÓN CON DETERMINADAS CONDUCTAS Y CON OTROS CONSTRUCTOS**.
2. **DISEÑAR EL INSTRUMENTO**, SELECCIONANDO **ÍTEMS RELEVANTES Y REPRESENTATIVOS** DEL CONSTRUCTO.
3. **OBTENER DATOS EMPÍRICOS** DE LAS RELACIONES ENTRE LAS **PUNTUACIONES DEL TEST Y LAS CONDUCTAS OBSERVABLES**.

ANÁLISIS FACTORIAL

ES QUIZÁS LA TÉCNICA MÁS UTILIZADA PARA **PONER A PRUEBA O DESCUBRIR LAS HIPÓTESIS PLANTEADAS ACERCA DE LA ESTRUCTURA INTERNA DEL CONSTRUCTO Y DE LAS RELACIONES DEL MISMO CON OTRAS VARIABLES**, NOS PERMITE, POR TANTO, DESCUBRIR LA ESTRUCTURA QUE SUBYACE A LAS PUNTUACIONES OBTENIDAS POR LOS SUJETOS EN LOS DISTINTOS ÍTEMS DEL TEST O EN UN CONJUNTO DE TEST.

EL AF SON UN CONJUNTO DE TÉCNICAS ESTADÍSTICAS EMPLEADAS PARA **REPRESENTAR Y EXPLICAR UN CONJUNTO DE VARIABLES OBSERVADAS EN UN MENOR NÚMERO DE VARIABLES LATENTES (NO OBSERVADAS) = CONSTRUCTOS**.

SIRVE PARA **VALIDAR LA ESTRUCTURA INTERNA DE UNA PRUEBA**. ANALIZA SI HAY UNA DIMENSIÓN O VARIAS EN EL TEST (CONSTRUCTOS/VARIABLES **UNIDIMENSIONALES** O **MULTIDIMENSIONALES**)



DOS ENFOQUES:

DESDE EL **ENFOQUE CONFIRMATIVO** SE ESTABLECE **HIPÓTESIS A PRIORI ACERCA DE LA ESTRUCTURA SUBYACENTE Y DEL NÚMERO DE DIMENSIONES** QUE SUBYACEN AL CONSTRUCTO, PERO NO PASA ASÍ EN EL **ENFOQUE EXPLORATORIO**.

- PARTIMOS DE n **MEDIDAS** (VARIABLES OBSERVABLES) TOMADAS A LA MISMA MUESTRA DE SUJETOS
- OBTENEMOS UNA **MATRIZ** ($n \times n$) CON LAS INTERCORRELACIONES ENTRE TODAS ELLAS.
- A TRAVÉS DEL **AF IDENTIFICAMOS LAS VARIABLES LATENTES O FACTORES**. OBTENEMOS:

- **VALIDEZ CONVERGENTE**: EN UN MISMO FACTOR SE AGRUPAN MÚLTIPLES INDICADORES DE CONSTRUCTO.
- **VALIDEZ DISCRIMINANTE**: SE OBTIENEN MEDIDAS DE OTROS CONSTRUCTOS Y SE AGRUPAN EN DISTINTOS FACTORES.

LOS MARCIANOS SE VAN DE JUERGA
(FUENTE: SERGIO ESCORIAL, UCM)

MATRIZ MULTIMÉTODO-MULTIRRASGO MMMR

PERMITE EVALUAR LA VALIDEZ CONVERGENTE Y DISCRIMINANTE DE LOS TEST Y ANALIZAR SU ESTRUCTURA EXTERNA. CONSISTE EN MEDIR UN MISMO CONSTRUCTO MEDIANTE DISTINTOS PROCEDIMIENTOS Y DISTINTOS CONSTRUCTOS MEDIANTE EL MISMO PROCEDIMIENTO, Y CONSTRUIR UNA **MATRIZ DE CORRELACIONES**.

• VALIDEZ CONVERGENTE (VC)

CORRELACIONES ENTRE LAS MEDIDAS OBTENIDAS DEL **MISMO CONSTRUCTO** A TRAVÉS DE **DISTINTOS PROCEDIMIENTOS (TEST)** **MISMO RASGO-DISTINTO MÉTODO**: MONORRASGO - HETEROMÉTODO. **EXISTE SIEMPRE QUE ESTAS CORRELACIONES SEAN ALTAS**.

• VALIDEZ DISCRIMINANTE(VD)

CORRELACIONES ENTRE LAS MEDIDAS OBTENIDAS DE **DISTINTOS CONSTRUCTOS** CON EL **MISMO PROCEDIMIENTO (TEST)** **DISTINTO RASGO-MISMO MÉTODO**: HETERORRASGO – MONOMÉTODO. **EXISTE SIEMPRE QUE LAS CORRELACIONES ANTERIORES SEAN SIGNIFICATIVAMENTE MÁS ALTAS QUE ÉSTAS**. $VD < VC$

- **COEFICIENTES DE FIABILIDAD**: CORRELACIONES QUE SE ENCUENTRAN EN LA DIAGONAL DE LA MATRIZ. **MISMO RASGO, MISMO MÉTODO**, HETERORRASGO, HETEROMÉTODO RELACIÓN ENTRE CADA CONSTRUCTO MEDIDO CON EL MISMO MÉTODO

INCONVENIENTE. NO HAY UN **CRITERIO ESTADÍSTICO** QUE PERMITA TOMAR DECISIONES. SÓLO SE PUEDE AFIRMAR QUE PARECE HABER EVIDENCIA. ACTUALMENTE PARA OBTENER MÁS INFORMACIÓN SE ESTÁ UTILIZANDO EL AFC.

		Extraversión			Liderazgo			Inteligencia social		
		AI	OS	EP	AI	OS	EP	AI	OS	EP
Extraversión	AI	0.89								
	OS	0.51	0.89							
	EP	0.38	0.37	0.76						
Liderazgo	AI	0.57	0.22	0.09	0.93					
	OS	0.22	0.57	0.10	0.68	0.94				
	EP	0.11	0.11	0.46	0.59	0.58	0.84			
Inteligencia social	AI	0.56	0.22	0.11	0.67	0.42	0.33	0.94		
	OS	0.23	0.58	0.12	0.43	0.66	0.34	0.67	0.92	
	EP	0.11	0.11	0.45	0.34	0.32	0.58	0.58	0.60	0.85

MATRIZ MULTIMÉTODO MULTIRRASGO – MMMR

MATRIZ DE CORRELACIONES EN LA QUE **3 RASGOS**: EXTRAVERSIÓN, LIDERAZGO E INTELIGENCIA SOCIAL SE MIDIERON CON **3 MÉTODOS** DISTINTOS: AI AUTOINFORME, OS OBSERVACIÓN SISTEMÁTICA Y EP ENCUESTA A PROFESORES.

3. VALIDEZ REFERIDA AL CRITERIO

EL OBJETIVO PRINCIPAL ES **EVALUAR LA HIPÓTESIS DE RELACIÓN ENTRE EL TEST Y CRITERIO.**

SE UTILIZAN **2 TIPOS DE ÍNDICES O MEDIDAS** PARA DESCRIBIR LA CAPACIDAD DE UN TEST PARA PREDECIR UN CRITERIO:

- **MEDIDAS CORRELACIONALES** (COEFICIENTE DE VALIDEZ, DE DETERMINACIÓN, DE ALIENACIÓN, DE VALOR PREDICTIVO)
- **MEDIDAS DE ERROR EN LA PREDICCIÓN** (ERRORES DE ESTIMACIÓN)

A DIFERENCIA DE LO QUE OCURRÍA EN EL PROCESO DE VALIDACIÓN DEL CONSTRUCTO, **SE ACENTÚA EL INTERÉS EN EL ASPECTO EMPÍRICO DEL PROCESO MÁS QUE EN EL TEÓRICO.**

PASOS PARA DISEÑAR UN ESTUDIO DE VALIDACIÓN REFERIDA AL CRITERIO

1. DEFINIR CLARAMENTE EL **CRITERIO** QUE SE QUIERE MEDIR
2. IDENTIFICAR EL **INDICADOR O INDICADORES** QUE SE VAN A UTILIZAR PARA LA MEDIDA DEL CRITERIO
3. SELECCIONAR UNA **MUESTRA** DE SUJETOS REPRESENTATIVA
4. **APLICAR EL TEST** A LA MUESTRA DE SUJETOS Y OBTENER UNA PUNTUACIÓN PARA CADA UNO
5. OBTENER UNA **MEDIDA DE CADA SUJETO EN EL CRITERIO** BIEN EN EL MISMO MOMENTO DE APLICACIÓN DEL TEST (V. **CONCURRENTES**) O BIEN AL CABO DE CIERTO TIEMPO (V. **PREDICTIVA**)
6. **DETERMINAR EL GRADO DE RELACIÓN** ENTRE LAS PUNTUACIONES OBTENIDAS EN EL TEST Y LA MEDIDA DEL CRITERIO

EL PROBLEMA DE LA SELECCIÓN Y MEDICIÓN DEL CRITERIO

TODOS LOS INDICADORES SON PARCIALES Y NO OFRECEN UNA COMPRENSIÓN COMPLETA DEL CRITERIO.

CRITERIO: ÉXITO COMO PROFESOR DE PSICOMETRÍA. **INDICADORES:** CONOCIMIENTO, PUBLICACIONES, PROYECTOS DE INVESTIGACIÓN, EMPATÍA CON LOS ALUMNOS, CAPACIDAD DE ORGANIZACIÓN,... **¿CÓMO ELEGIR CUÁL ES EL INDICADOR MÁS ADECUADO?**

ELECCIÓN DE LOS INDICADORES (THORNDIKE Y HAGEN, 1989):

- QUE SEAN **RELEVANTES** (JUICIOS DE EXPERTOS, CRITERIOS RACIONALES)
- QUE ESTÉN **LIBRES DE SESGOS**
- QUE SEAN **FIABLES** (→ ESTABLES)
- DEBEN SER **ACCESIBLES** (NO PROBLEMAS ECONÓMICOS, TIEMPO PARA OBTENER SU MEDIDA,...)

PROCEDIMIENTOS ESTADÍSTICOS UTILIZADOS EN LA VALIDACIÓN REFERIDA AL CRITERIO

4 TIPOS DE DISEÑO

- UN **ÚNICO TEST PREDICTOR** Y UN **INDICADOR DEL CRITERIO**
CORRELACIÓN (GRADO DE RELACIÓN) Y **MODELO DE REGRESIÓN LINEAL SIMPLE** (PRONOSTICAR) (**TEMA 6**)
- **VARIOS PREDICTORES** Y UN **INDICADOR DEL CRITERIO**
CORRELACIÓN Y **MODELO DE REGRESIÓN LINEAL MÚLTIPLE** (**TEMA 7**)
- **VARIOS PREDICTORES** Y **VARIOS INDICADORES DEL CRITERIO**, CUANTITATIVOS AMBOS
CORRELACIÓN CANÓNICA Y **MODELO DE REGRESIÓN LINEAL MULTIVARIANTE**
- **PROCEDIMIENTOS BASADOS EN LA TEORÍA DE DECISIÓN:** VALIDEZ Y UTILIDAD DE LAS DECISIONES
MÉTODOS PARA OPTIMIZAR LAS DECISIONES: MAXIMIN, MINIMAX, TEORÍA DE UTILIDAD MULTIATRIBUTO

4. VALIDACIÓN CON UN SOLO PREDICTOR Y UN SOLO INDICADOR DE CRITERIO

EL COEFICIENTE DE VALIDEZ

EL TIPO DE CORRELACIÓN ENTRE TEST (X) Y CRITERIO (Y) DEPENDERÁ DE LA **NATURALEZA DE LAS VARIABLES** IMPLICADAS (NIVEL DE MEDIDA DEL TEST Y DEL CRITERIO)

CONTINUA

DICOTÓMICA (SE PRESENTA EN DOS ALTERNATIVAS DE MANERA NATURAL)

DICOTOMIZADA (LA CONVERTIMOS* EN DOS)

SE USA LA TABLA 6 DEL FORMULARIO



$$r_{xy} = \frac{n \sum XY - \sum X \sum Y}{\sqrt{n \sum X^2 - (\sum X)^2} \sqrt{n \sum Y^2 - (\sum Y)^2}}$$

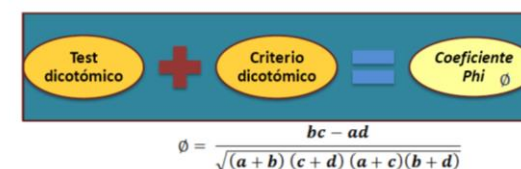
$$r_{bp} = \frac{\bar{X}_p - \bar{X}}{S_x} \sqrt{p/q}$$



$$\phi_b = \frac{bc - ad}{\sqrt{(a+b)(c+d)}} \sqrt{\frac{p}{q}}$$



$$r_b = \frac{\bar{X}_p - \bar{X}}{S_x} p/y$$



$$\phi = \frac{bc - ad}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$$

		X
	0	1
Y	0	a b
	1	c d



EL MODELO DE REGRESIÓN LINEAL (VER HOJAS CENTRO ASOCIADO A CORUÑA)

UNA VEZ CONOCIDO EL GRADO DE RELACIÓN ENTRE TEST Y CRITERIO SE PUEDEN HACER PRONÓSTICOS. SIRVEN **PARA PRONOSTICAR O PREDECIR COMO SE COMPORTARÁ LA VARIABLE CRITERIO, TOMANDO EN CONSIDERACIÓN LA VARIABLE PREDICTORA X**. ES DECIR, PARA HACER ESTIMACIONES DE LA PUNTUACIÓN DEL CRITERIO UNA VEZ OBTENIDAS LAS PUNTUACIONES DE LOS SUJETOS EN UN TEST.

PUNTUACIÓN	RECTA REGRESIÓN	SIGNIFICADO	CÁLCULO
DIRECTAS	X $Y' = A + B X$	A ordenada en el origen o intercepto B pendiente, coeficiente de regresión o tasa de cambio	$Y' = r_{xy} \frac{S_y}{S_x} (X - \bar{X}) + \bar{Y}$
DIFERENCIALES	$x = X - \bar{X}$ $y' = b x$	$b = B$ pendiente	$y' = r_{xy} \frac{S_y}{S_x} (X - \bar{X})$
TÍPICAS	$Z_x = \frac{X - \bar{X}}{S_x}$ $Z_{y'} = r_{xy} Z_x$	r_{xy} pendiente en típicas	$r_{xy} = \frac{S_{xy}}{S_x S_y}$

	DIRECTAS	DIFERENCIALES	TÍPICAS
VARIANZA TOTAL		$S_y^2 = S_{y'}^2 + S_{y.x}^2$	
ERROR TÍPICO DE MEDIDA	$S_{y.x} = S_y \sqrt{1 - r_x^2}$	$S_{y.x} = S_y \sqrt{1 - r_x^2}$	$S_{zy.x} = \sqrt{1 - r_x^2}$
ERROR MÁXIMO	$E_{max} = Z_{\alpha/2} S_{y.x}$	$E_{max} = Z_{\alpha/2} S_{y.x}$	$E_{max} = Z_{\alpha/2} S_{zy.x}$
INTERVALO CONFIDENCIAL	$Y' \pm E_{max}$	$y' \pm E_{max}$	$Z_{y'} \pm E_{max}$

INTERPRETACIÓN DE LA EVIDENCIA OBTENIDA ACERCA DE LA CAPACIDAD PREDICTORA DEL TEST

COEFICIENTES	SIGNIFICADO	CÁLCULO
DE DETERMINACIÓN $C.D. = D = r_{xy}^2$ COEFICIENTE DE VALIDEZ AL CUADRADO	PROPORCIÓN DE LA VARIANZA DE LAS PUNTUACIONES DE LOS SUJETOS EN EL CRITERIO (VARIABLE DEPENDIENTE) QUE SE PUEDE PRONOSTICAR A PARTIR DEL TEST (VARIABLE INDEPENDIENTE) <i>*ES EL MÁS IMPORTANTE: PORCENTAJE QUE SE DEBE A LA CAUSA.</i> PROPORCIÓN DE LA VARIANZA DE LAS PUNTUACIONES DE LOS SUJETOS EN EL CRITERIO QUE NO SE PUDE PRONOSTICAR A PARTIR DEL TEST X	$r_{xy}^2 = \frac{S_{y'}^2}{S_y^2} = 1 - \frac{S_{yx}^2}{S_y^2}$ $1 - r_{xy}^2 = \frac{S_{yx}^2}{S_y^2}$
DE ALIENACIÓN $C.A. = K = \sqrt{1 - r_{xy}^2}$	PROPORCIÓN DEL ERROR TÍPICO DE ESTIMACIÓN RESPECTO A LA DESVIACIÓN TÍPICA DE LAS PUNTUACIONES EN EL CRITERIO. REPRESENTA LA INSEGURIDAD, O EL AZAR, QUE AFECTA A LOS PRONÓSTICOS.	
VALOR PREDICTIVO $C.V.P. = E = 1 - \sqrt{1 - r_{xy}^2}$ $C.V.P. = 1 - C.A.$	CAPACIDAD DEL TEST PARA PRONOSTICAR EL CRITERIO PROPORCIÓN DE SEGURIDAD EN LOS PRONÓSTICOS	

TEMA 7. VALIDEZ DE LAS INFERENCIAS II

ÍNDICES DE VALIDEZ Y DE SELECCIÓN

PROCEDIMIENTOS PARA ANALIZAR LA VALIDEZ DE LAS DECISIONES TOMADAS A PARTIR DE LAS PUNTUACIONES OBTENIDAS POR LOS SUJETOS EN UN TEST (O VARIOS) EN RELACIÓN A UN CRITERIO DICOTÓMICO. ANALIZAMOS SI LA CLASIFICACIÓN REALIZADA POR EL TEST COINCIDE CON LA REALIZADA POR JUECES EXTERNOS EN EL CRITERIO. LAS PUNTUACIONES ESTÁN DICOTOMIZADAS A PARTIR DE UN PUNTO DE CORTE DE MANERA QUE PERMITA ASIGNAR A LOS SUJETOS EN LAS DOS CATEGORÍAS: ADMITIDOS/RECHAZADOS, APTOS/NO APTOS, ENFERMOS/NO ENFERMOS. CUANDO LAS DECISIONES TOMADAS A PARTIR DEL TEST (X) SE VEN CONFIRMADAS CON LAS DECISIONES TOMADAS EN EL CRITERIO (Y), ENTONCES LA PRUEBA ES VÁLIDA.		<table border="1"> <thead> <tr> <th colspan="2" rowspan="2"></th><th colspan="2">CRITERIO</th><th rowspan="2"></th></tr> <tr> <th>A</th><th>R</th></tr> </thead> <tbody> <tr> <th rowspan="2">TEST</th><th>A</th><td>Aciertos N_{AA}</td><td>Falsos Positivos N_{AR}</td><td>N_{TA}</td></tr> <tr> <th>R</th><td>Falsos Negativos N_{RA}</td><td>Aciertos N_{RR}</td><td>N_{TR}</td></tr> <tr> <td></td><td></td><td>N_{AC}</td><td>N_{RC}</td><td>N</td></tr> </tbody> </table> N_{AC} ALUMNOS CONSIDERADOS APTOS EN EL CRITERIO. N_{RC} ALUMNOS CONSIDERADOS NO APTOS EN EL CRITERIO N_{AT} ALUMNOS CONSIDERADOS APTOS EN EL TEST N_{RT} ALUMNOS CONSIDERADOS NO APTOS EN EL TEST				CRITERIO			A	R	TEST	A	Aciertos N_{AA}	Falsos Positivos N_{AR}	N_{TA}	R	Falsos Negativos N_{RA}	Aciertos N_{RR}	N_{TR}			N_{AC}	N_{RC}	N	$N_{AA} + N_{RR} = \text{Aciertos}$ ALUMNOS QUE HAN SIDO CLASIFICADOS DEL MISMO MODO EN LA PRUEBA DE SELECCIÓN (TEST) Y EN EL CRITERIO. $N_{RA} = \text{Falsos Negativos}$ ALUMNOS QUE SUPERARON EL CRITERIO DE RENDIMIENTO Y SIN EMBARGO EN LA PRUEBA DE ADMISIÓN NO SUPERARON EL PUNTO DE CORTE. EN UN PROCESO DE SELECCIÓN HABRÍAN SIDO RECHAZADOS Y, SIN EMBARGO, DEBERÍAN HABER SIDO ADMITIDOS. $N_{AR} = \text{Falsos Positivos}$ ALUMNOS QUE EN LA PRUEBA DE ADMISIÓN SUPERARON EL PUNTO DE CORTE Y LUEGO NO SUPERARON EL CRITERIO DE RENDIMIENTO. EN UN PROCESO DE SELECCIÓN NO DEBERÍAN HABER SIDO SELECCIONADOS Y, SIN EMBARGO, AL SUPERAR EL PUNTO DE CORTE EN EL PREDICTOR, SERÍAN ADMITIDOS.
		CRITERIO																							
		A	R																						
TEST	A	Aciertos N_{AA}	Falsos Positivos N_{AR}	N_{TA}																					
	R	Falsos Negativos N_{RA}	Aciertos N_{RR}	N_{TR}																					
		N_{AC}	N_{RC}	N																					
ÍNDICES DE VALIDEZ DIRIGIDOS A ESTUDIAR LA COINCIDENCIA EN LAS CLASIFICACIONES REALIZADAS A PARTIR DEL TEST Y LA REALIZADA EN EL CRITERIO POR LOS JUECES. ESTOS PROCEDIMIENTOS SON LOS QUE SE UTILIZAN GENERALMENTE EN LOS TRC	COEFICIENTE DE KAPPA	PERMITE EVALUAR LA CONSISTENCIA O ACUERDO ENTRE LAS DECISIONES ADOPTADAS A PARTIR DE LAS PUNTUACIONES OBTENIDAS POR LOS SUJETOS EN EL PREDICTOR Y EN EL CRITERIO.	$K = \frac{F_c - F_d}{N - F_d}$ $F_c = N_{AA} + N_{RR}$ $F_d = N_{TA} + N_{TR}$	$Nt_{AA} = \frac{N_{AC} * N_{TA}}{N}$ $Nt_{RR} = \frac{N_{RC} * N_{RT}}{N}$																					
	PROPORCIÓN DE CLASIFICACIONES CORRECTAS	PROPORCIÓN DE CLASIFICACIONES CORRECTAS TANTO EN EL TEST COMO EN EL CRITERIO.		$P_{CC} = \frac{N_{AA} + N_{RR}}{N}$																					
	SENSIBILIDAD	PROPORCIÓN DE ASPIRANTES CORRECTA/ SELECCIONADOS MEDIANTE LA PRUEBA DE ADMISIÓN RESPECTO DE TODOS QUE TUVIERON ÉXITO EN EL CRITERIO		$S = \frac{N_{AA}}{N_{AC}}$																					
	ESPECIFICIDAD	PROPORCIÓN DE ASPIRANTES QUE FUERON CORRECTA/ RECHAZADOS MEDIANTE LA PRUEBA DE ADMISIÓN.		$E = \frac{N_{RR}}{N_{RC}}$																					
ÍNDICES DE SELECCIÓN	RAZÓN DE EFICACIA	PROPORCIÓN DE ASPIRANTES SELECCIONADOS MEDIANTE LA PRUEBA DE ADMISIÓN.		$RE = \frac{N_{AA}}{N_{TA}}$																					
	RAZÓN DE IDONEIDAD	PROPORCIÓN DE ASPIRANTES QUE RINDIERON SATISFACTORIAMENTE EL CRITERIO .		$RI = \frac{N_{AC}}{N}$																					
	RAZÓN DE SELECCIÓN	PROPORCIÓN DE ASPIRANTES QUE HAN SIDO SELECCIONADOS MEDIANTE EL TEST .		$RS = \frac{N_{TA}}{N}$																					

¿DÓNDE SITUAR LOS PUNTOS DE CORTE?

<table border="1"> <thead> <tr> <th colspan="2" rowspan="2"></th><th colspan="2">CRITERIO</th></tr> <tr> <th>A</th><th>R</th></tr> </thead> <tbody> <tr> <th rowspan="2">TEST</th><th>A</th><td>Aciertos</td><td>Falsos Positivos</td></tr> <tr> <th>R</th><td>Falsos Negativos</td><td>Aciertos</td></tr> </tbody> </table>				CRITERIO		A	R	TEST	A	Aciertos	Falsos Positivos	R	Falsos Negativos	Aciertos	DOS CRITERIOS A UTILIZAR EN LA TOMA DE DECISIONES EN AMBIENTE DE INCERTIDUMBRE: CRITERIO MAXIMIN (MÁXIMO DE LOS MÍNIMOS) SE ELEGIRÁ AQUELLA ALTERNATIVA QUE ENTRE LOS RESULTADOS MÁS DESFAVORABLES, LE PERMITA OBTENER LA MÁXIMA GANANCIA. CRITERIO MINIMAX (MÍNIMO DE LOS MÁXIMOS) ENTRE LAS ALTERNATIVAS DE UNA MAYOR PERDIDA SE ELEGIRÁ AQUELLA QUE PROPORCIONARÁ UNA «PÉRDIDA» MENOR.	CRITERIO DE SELECCIÓN MÁS Estricto: EL PUNTO DE CORTE POR EL TEST SE INCREMENTA. DISMINUYE N_{AR} (FALSOS POSITIVOS) HABRÍA UN MEJOR NUMERO DE ASPIRANTES QUE HABIENDO SIDO SELECCIONADOS POR EL TEST NO CUMPLIERAN EL RENDIMIENTO ADECUADO EN EL CRITERIO. AUMENTA N_{RA} (FALSOS NEGATIVOS) SE QUEDARÍAN FUERA UN MAYOR NÚMERO DE ASPIRANTES QUE, SI HUBIERAN SIDO SELECCIONADOS, PODRÍAN HABER RENDIDO ADECUADAMENTE EN EL CRITERIO. CRITERIO MÁS Estricto: EL PUNTO DE CORTE DEL CRITERIO SE INCREMENTA. DISMINUYE LOS FALSOS NEGATIVOS. AUMENTA LOS FALSOS POSITIVOS.
				CRITERIO												
		A	R													
TEST	A	Aciertos	Falsos Positivos													
	R	Falsos Negativos	Aciertos													

MODELOS DE SELECCIÓN

COMPENSATORIO	SE TRATA DE UN MODELO ADITIVO EN EL QUE A CADA SUJETO SE LE ASIGNA UNA <u>ÚNICA PUNTUACIÓN GLOBAL</u> . LOS SUJETOS PUEDEN COMPENSAR UNA BAJA PUNTUACIÓN EN UNA DE LAS PRUEBAS CON UNA PUNTUACIÓN ALTA EN OTRAS, DE MANERA QUE EL RESULTADO FINAL SEA UNA ÚNICA PUNTUACIÓN OBTENIDA UTILIZANDO EL <u>MODELO REGRESIÓN LINEAL MÚLTIPLE</u> .
CONJUNTIVO	SE FIJAN DE ANTEMANO UNOS MÍNIMOS Y SÓLO SE SELECCIONAN AQUELLAS PERSONAS QUE HAYAN <u>SUPERADO ESOS MÍNIMOS EN TODAS Y CADA UNA DE LAS PRUEBAS</u> .
DISYUNTIVO	SOLO SE EXIGE SUPERAR <u>DETERMINADO NIVEL DE COMPETENCIA EN AL MENOS ALGUNO DE LOS PREDICTORES</u> O BLOQUE DE PREDICTORES.
CONJUNTIVO-COMPENSATORIO	PRIMERO SE SELECCIONA AQUELLOS SUJETOS QUE SUPERAR LOS <u>MÍNIMOS</u> ESTABLECIDOS (<u>MODELO CONJUNTIVO</u>) Y LUEGO SE ORDENA EN FUNCIÓN DE LA <u>PUNTUACIÓN GLOBAL OBTENIDA</u> (MODELO COMPENSATORIO)
DISYUNTIVO-COMPENSATORIO	SE HACE UNA PRIMERA <u>SELECCIÓN</u> APLICANDO EL <u>MODELO DISYUNTIVO</u> Y A LOS SUJETOS SELECCIONADOS SE LES APLICA EL <u>MODELO COMPENSATORIO</u> .

ESTIMAR LA EFICACIA DE UNA SELECCIÓN

1. A TRAVÉS DE LA <u>RAZÓN DE EFICACIA</u> : PROPORCIÓN DE PERSONAS SELECCIONADAS QUE TIENEN ÉXITO EN EL CRITERIO	$RE = \frac{N_{AA}}{N_{TA}}$	2. A TRAVÉS DEL MODELO DE REGRESIÓN PUES PERMITE ESTIMAR LA <u>PROBABILIDAD DE QUE LOS SELECCIONADOS TENGAN ÉXITO EN EL CRITERIO</u> .	$Z_c = \frac{Y_c - Y'}{S_{y \cdot c}}$	$P(Z \leq Z_c)$ $P(Z \geq Z_c) = 1 - P(Z \leq Z_c)$
---	------------------------------	--	--	--

3. FACTORES QUE INFLUYEN EN EL COEFICIENTE DE VALIDEZ

VARIABILIDAD DE LA MUESTRA	LONGITUD	FIABILIDAD DE LAS PUNTUACIONES DEL TEST Y DEL CRITERIO:			
LA VALIDEZ r_{xy} TIENDE A <u>AUMENTAR A MEDIDA QUE LA VARIABILIDAD DE LA MUESTRA UTILIZADA ES MAYOR</u> . DOS GRUPOS: <u>ASPIRANTES</u> (MAYÚSCULAS) Y <u>SELECCIONADOS</u> (MINÚSCULAS) SUPUESTOS LAS <u>PENDIENTES</u> DE LOS DOS GRUPOS SON IGUALES: $B = b \Rightarrow R_{XY} \frac{S_Y}{S_X} = r_{xy} \frac{S_Y}{S_X}$ COEFICIENTE DE VALIDEZ DE LA BATERÍA: $R_{xy} = \frac{S_X r_{xy}}{\sqrt{S_X^2 r_{xy}^2 + S_X^2 - S_X^2 r_{xy}^2}}$	<u>AL AUMENTAR LA LONGITUD DEL TEST AUMENTA LA FIABILIDAD</u> , Y AL MEJORAR ÉSTA AUMENTA LA VALIDEZ DEL TEST. $R_{xy} = \frac{r_{xy} \sqrt{n}}{\sqrt{1 + (n-1)r_{xx}}}$ $n = \frac{\text{Finales}}{\text{Iniciales}}$ $n = \frac{R_{xy}^2 (1 - r_{xx})}{r_{xy}^2 - R_{xy}^2 r_{xx}}$	SPEARMAN PROPUSO LA <u>FÓRMULA DE ATENUACIÓN</u> QUE PERMITE <u>CORREGIR LA ATENUACIÓN, DISMINUCIÓN O REDUCCIÓN DEL COEFICIENTE DE VALIDEZ</u> DEBIDA A LA PRESENCIA DE LOS ERRORES DE MEDIDA: EL <u>VALOR MÁXIMO DEL COEFICIENTE DE VALIDEZ</u> ES EL <u>ÍNDICE DE FIABILIDAD</u> : $r_{xy} \leq r_{xv}; r_{xy} \leq \sqrt{r_{xx}}$			
		SUPUESTOS	TEST Y CRITERIO	TEST	CRITERIO
		MEJORA	$R_{xy} = \frac{r_{xy}}{\sqrt{\frac{r_{xx}}{R_{xx}}} \sqrt{\frac{r_{yy}}{R_{yy}}}}$	$R_{xy} = \frac{r_{xy}}{\sqrt{\frac{r_{xx}}{R_{xx}}}}$	$R_{xy} = \frac{r_{xy}}{\sqrt{\frac{r_{yy}}{R_{yy}}}}$
		FIABILIDAD PERFECTA	$R_{vxy} = \frac{r_{xy}}{\sqrt{r_{xx} r_{yy}}}$	$R_{vxy} = \frac{r_{xy}}{\sqrt{r_{xx}}}$	$R_{vxy} = \frac{r_{xy}}{\sqrt{r_{yy}}}$

TEMA 8. ANÁLISIS DE LOS ÍTEMS

PROPIEDADES PSICOMÉTRICAS DE LOS ÍTEMS

EL ANÁLISIS DE LOS ELEMENTOS DE UN TEST SE PUEDE LLEVAR A CABO MEDIANTE **DOS PROCEDIMIENTOS COMPLEMENTARIOS**:

- ANÁLISIS CUANTITATIVO** (OBTENCIÓN DE **ESTADÍSTICOS** BASADOS EN LA DISTRIBUCIÓN DE LAS RESPUESTAS DE LOS SUJETOS A LAS DISTINTAS ALTERNATIVAS) PASOS: (1) **ESTIMAR SUS PARÁMETROS**: DIFICULTAD, DISCRIMINACIÓN, FIABILIDAD Y VALIDEZ, (2) **ANÁLISIS DE LOS DISTRACTORES**, Y (3) **SESGO**: FUNCIONAMIENTO DIFERENCIAL DE LOS ÍTEMS (FDI)
- ANÁLISIS CUALITATIVO: JUICIOS DE EXPERTOS** (EVALUAR LA CALIDAD MÉTRICA DEL ÍTEM EN FUNCIÓN DE CRITERIOS CONOCIDOS) LA **VALIDEZ DE CONTENIDO** SE LLEVA A CABO A TRAVÉS DE **ÍNDICES DE ACUERDO O CONGRUENCIA** ENTRE JUECES, POR EJEMPLO, EL **CRV** (PROMEDIO QUE TOMA COMO REFERENCIA EL NÚMERO DE JUECES QUE CONSIDERAN UN ÍTEM “NO NECESARIO”, “ÚTIL” O “ESENCIAL”) Y EL **ÍNDICE DE CONGRUENCIA DE ROVINELLI Y HAMBLETON** (SE PIDE A CADA UNO DE LOS JUECES QUE VALOREN EN UNA ESCALA DE 3 PUNTOS (-1, 0, 1) O 5, EL GRADO EN QUE EL ÍTEM ESTÁ RELACIONADO CON LA DIMENSIÓN QUE TRATA DE MEDIR) EL OBJETIVO ES **IDENTIFICAR Y REDUCIR LAS FUENTES DE ERROR**, TANTO ALEATORIO COMO SISTEMÁTICO PARA PODER **ELIMINAR** AQUELLOS QUE NO REÚNEN LAS SUFICIENTES GARANTÍAS PSICOMÉTRICAS.

ESTE TEMA VAMOS A ABORDAR EL **ANÁLISIS CUANTITATIVO O ESTADÍSTICO DE LOS ÍTEMS**, PUESTO QUE EL ANÁLISIS CUALITATIVO YA SE HA ABORDADO A LO LARGO DE LOS TEMAS PREDECENTES. EL ANÁLISIS SE LLEVA A CABO EN DOS FASES: EN EL ESTUDIO PILOTO Y EN LA CONFECCIÓN DE LA PRUEBA FINAL.

NOS CENTRAMOS EN LAS CARACTERÍSTICAS QUE **VAN A AFECTAR A LA CALIDAD GLOBAL DEL TEST**:

PROPIEDADES PSICOMÉTRICAS DE LOS ÍTEMS	LA DIFICULTAD	ASPECTO QUE SE ESTABLECE A PRIORI DESDE CONSIDERACIONES TEÓRICAS Y QUE ESTÁ EN FUNCIÓN DEL CONTENIDO, PERO QUE TIENE QUE CONTRASTARSE A NIVEL EMPÍRICO (PORCENTAJE DE SUJETOS QUE LO RESPONDEN CORRECTAMENTE)	RELACIONADAS CON LA MEDIA Y LA VARIANZA DE LA PUNTUACIÓN TOTAL DEL TEST.
	EL PODER DISCRIMINANTE	CAPACIDAD DE DIFERENCIAR ADECUADAMENTE A LOS SUJETOS DE DIFERENTES NIVELES DE UNA VARIABLE CRITERIO.	
	LA FIABILIDAD Y LA VALIDEZ	RELACIONADAS CON LA DESVIACIÓN TÍPICA DEL TEST. INDICAN LA CONTRIBUCIÓN DEL ÍTEM A LA FIABILIDAD Y VALIDEZ TOTALES .	
	COMPARACIÓN DE LAS RESPUESTAS DE LOS ÍTEMS	MEDIR EL GRADO DE ACUERDO U HOMOGENEIDAD ENTRE LOS ÍTEMS, PARA VER SI MIDEN LA MISMA VARIABLE.	
	ANÁLISIS DE DISTRACTORES	APORTA EVIDENCIAS SOBRE LA UTILIDAD DE CADA ALTERNATIVA Y SU CONTRIBUCIÓN A LA CALIDAD MÉTRICA DEL ÍTEM.	
	FUNCIONAMIENTO DIFERENCIAL DEL ÍTEM	ANALIZA SI DE MANERA SISTEMÁTICA SUJETOS DE DISTINTOS GRUPOS DE PERTENENCIA PERO CON EL MISMO NIVEL EN EL RASGO MEDIDO TIENEN DISTINTAS PROBABILIDADES DE ÉXITO EN UN ÍTEM , PERO SIN APUNTAR A POSIBLES CAUSAS.	
LOS ÍTEMS PUEDEN ADOPTAR DISTINTOS FORMATOS Y EVALUAR VARIABLES COGNITIVAS (APTITUDES, RENDIMIENTO,...) Y VARIABLES NO COGNITIVAS (ACTITUDES, INTERESES, VALORES,...) LOS ESTADÍSTICOS QUE PRESENTAMOS SE UTILIZAN, FUNDAMENTALMENTE, CON ÍTEMS APTITUDINALES O DE RENDIMIENTO EN LOS QUE EXISTE RESPUESTA CORRECTA Y UNA O VARIAS INCORRECTAS . PARA LLEVAR A CABO EL ANÁLISIS LO PRIMERO ES DISPONER DE UNA MATRIZ DE DATOS CON LAS RESPUESTAS DE LOS SUJETOS A CADA UNO DE LOS ÍTEMS. PARA EL ANÁLISIS DE PUNTUACIONES DEL TEST Y DE RESPUESTAS A LA ALTERNATIVA CORRECTA , LA MATRIZ SERÁ DE 1/0 (ACIERTOS/FALLOS) Y PARA EL DE RESPUESTAS A LAS ALTERNATIVAS INCORRECTAS DEBERÁN APARECER LAS OPCIONES CONCRETAS QUE SE HAYAN SELECCIONADO.			

1. DIFICULTAD DE UN ÍTEM

SOLO PARA TESTS QUE MIDEN VARIABLES COGNITIVAS, EN LOS QUE EXISTEN RESPUESTAS CORRECTAS E INCORRECTAS (**ÍTEMS DICOTÓMICOS O DICOMOTIZADOS**) **PROPORCIÓN DE SUJETOS QUE HAN RESPONDIDO CORRECTAMENTE** AL MISMO.

$$ID = \frac{A}{N}, \quad 0 < ID < 1$$

A = NÚMERO DE SUJETOS QUE RESPONDEN CORRECTAMENTE AL ÍTEM
N = NÚMERO DE SUJETOS QUE RESPONDEN AL ÍTEM

LA DIFICULTAD ES DEPENDIENTE DE LA MUESTRA DE SUJETOS UTILIZADA (DE TAL FORMA QUE UN MISMO ÍTEM TENDRÁ TANTOS ÍNDICES DE DIFICULTAD COMO MUESTRAS SE PUEDAN UTILIZAR PARA SU CÁLCULO)

- SE DEBERÍA DE LLAMAR **ÍNDICE DE FACILIDAD** YA QUE LOS VALORES ALTOS INDICAN LA FACILIDAD DEL ÍTEM
- ÍTEMS CON **VALORES EXTREMOS DEBEN SER DESCARTADOS** **ID = 0**, **ID = 1** YA QUE NO CONTRIBUYEN A DIFERENCIAR SUJETOS CON DISTINTO NIVEL EN EL RASGO. COMO NORMA GENERAL, EN LOS TEST DE APTITUDES SE CONSIGUEN MEJORES RESULTADOS CON ÍTEMS DE DIFICULTAD MEDIA. LÓGICAMENTE HABRÁ QUE INCLUIR ALGUNOS FÁCILES (AL PPIO) Y ALGUNO DIFÍCIL PARA DETECTAR A LOS DE PEOR NIVEL Y A LOS MEJORES.

LA DIFICULTAD DE LOS ÍTEMS AFECTA CLARAMENTE A LA VARIABILIDAD DE LAS PUNTUACIONES EN EL TEST Y A LA PRECISIÓN CON LA QUE LAS PUNTUACIONES DEL TEST DISCRIMINAN ENTRE DIFERENTES GRUPOS DE SUJETOS.

RELACIÓN ENTRE DIFICULTAD Y VARIANZA

EL ID ESTÁ DIRECTAMENTE RELACIONADO CON LA MEDIA Y LA VARIANZA DEL TEST. **LA SUMA DE TODAS LAS PUNTUACIONES OBTENIDAS POR LOS SUJETOS EN ESE ÍTEM ES IGUAL A LOS ACIERTOS**:

$$\sum X_j = A \Rightarrow \bar{X} = \sum_{j=1}^n ID_j$$

$$\mu = ID$$

EN ÍTEMS DICOTÓMICOS LA VARIANZA VIENE DADA POR:

$$S_j^2 = p_j q_j; \quad p_j = ID; \quad q_j = 1 - p_j$$

$$S_j^2 = ID(1 - ID)$$

UNA CUESTIÓN MUY RELEVANTE ES ENCONTRAR EL **VALOR DE p_j QUE MAXIMIZA LA VARIANZA**. ESTE VALOR ES **0.5**. ÍTEMS FALLADOS O ACERTADOS POR TODOS LOS SUJETOS REPRESENTAN UNA VARIANZA IGUAL A CERO.

CORRECCIÓN DE LOS ACIERTOS POR AZAR

SE APLICA **PARA EVITAR QUE SE ACIERTEN ÍTEMS POR AZAR, SIN CONOCER REALMENTE LA RESPUESTA**. CUÁNTO MAYOR SEA EL NÚMERO DE DISTRACTORES MENOS PROBABLE ES QUE LOS SUJETOS ACIERTEN EL ÍTEM POR AZAR. SE CONSIDERA QUE **TODAS LAS ALTERNATIVAS TIENEN LA MISMA PROBABILIDAD DE SER ELEGIDAS POR AZAR**.

$$ID_c = \frac{A}{N} - \frac{\frac{E}{k-1}}{N} = p - \frac{q}{k-1}$$

p PROPORCIÓN DE ACIERTOS, q PROPORCIÓN DE ERRORES

2. DISCRIMINACIÓN DE UN ÍTEM

- **CAPACIDAD DEL ÍTEM PARA DISTINGUIR O DIFERENCIAR A LOS SUJETOS DE BAJA, MEDIA Y ALTA PUNTUACIÓN EN UN CRITERIO.**

• SE DEBE CONOCER TANTO EN PRUEBAS COGNITIVAS COMO EN LAS QUE NO LO SON.

• EL CRITERIO PUEDE SER INTERNO (LA PUNTUACIÓN EN EL TEST) O EXTERNO (OTRO TEST U OTRA VARIABLE)

- CUANDO EL **CRITERIO ES INTERNO**, EL ÍNDICE DE DISCRIMINACIÓN SE PUEDE INTERPRETAR COMO **ÍNDICE DE HOMOGENEIDAD IH** PORQUE EXPRESA EL **GRADO DE SEMEJANZA ENTRE LAS RESPUESTA AL ÍTEM Y EL RESTO DE LOS ÍTEMS** MEDIADOS A TRAVÉS DE LA PUNTUACIÓN TOTAL EN EL TEST.
- CUANDO EL **CRITERIO ES EXTERNO** EL ÍNDICE DE DISCRIMINACIÓN ES UNA MEDIDA DEL **GRADO DE VALIDEZ DEL ÍTEM REFERIDO A ESE CRITERIO Y SU VALOR ES PROPORCIONAL AL ÍNDICE DE VALIDEZ IV** DEL ÍTEM.

DADO UN ÍTEM, LOS SUJETOS CON BUENAS PUNTUACIONES EN EL TEST HAN DE ACERTARLO EN MAYOR PROPORCIÓN QUE LOS QUE TIENEN BAJAS PUNTUACIONES. SI UN ÍTEM NO SIRVE PARA DIFERENCIAR DEBERÍA ELIMINARSE. YA HEMOS VISTO QUE **CUANDO TODAS LAS PERSONAS ACIERTAN EL ÍTEM ID = 1 O CUANDO NADIE LO ACIERTA ID = 0**, EL ÍTEM NO ES CAPAZ DE DISCRIMINAR. LA **MÁXIMA DISCRIMINACIÓN** TIENE LUGAR CUANDO **ID = 0.5** Y ES **ACERTADO POR EL 50%** DE LAS PERSONAS.

UNA PRIMERA APROXIMACIÓN AL CÁLCULO DE LA DISCRIMINACIÓN SERÍA CONTRASTAR LA **PROPORCIÓN DE ACIERTOS ENTRE DOS GRUPOS EXTREMOS DE APTITUD, UNO BAJO Y OTRO ALTO**. SI ES SUPERIOR EN EL ALTO O LA CORRELACIÓN ENTRE LAS PUNTUACIONES OBTENIDAS POR LOS SUJETOS EN EL ÍTEM Y EN EL TEST TOTAL ES POSITIVA, EL ÍTEM TENDRÍA PODER DISCRIMINATIVO. EN BASE A ESTO TENEMOS VARIAS FORMAS DE ESTUDIAR LA DISCRIMINABILIDAD DEL ÍTEM.

ÍNDICE DE DISCRIMINACIÓN BASADO EN GRUPOS EXTREMOS -1 < D < 1	BASADA EN LA COMPARACIÓN DEL RENDIMIENTO DE LOS DOS GRUPOS EXTREMOS (BAJO Y ALTO) • SE FORMAN LOS GRUPOS EXTREMOS DE BAJA Y ALTA PUNTUACIÓN <u>CON EL 27% (O 25%) INFERIOR Y SUPERIOR</u> DE LOS SUJETOS. LUEGO SE CALCULA D. • SI D = 1 TODAS LAS PERSONAS DEL GRUPO SUPERIOR HAN ACERTADO EL ÍTEM, Y LAS DEL INFERIOR FALLADO. <u>VALOR IDEAL</u>. • SI D = 0 EL ÍTEM <u>NO DISCRIMINA ADECUADAMENTE</u> ENTRE LOS SUJETOS DE DIFERENTES NIVEL DE APTITUD. • SI D < 0 EL ÍTEM NO CONTRIBUYE A MEDIR LA VARIABLE DE INTERÉS Y ADEMÁS <u>FAVORECE LA DISMINUCIÓN DE LA PRECISIÓN DEL INSTRUMENTO DE MEDIDA</u>. LOS MENOS COMPETENTES ACIERTAN MÁS EL ÍTEM QUE LOS COMPETENTES. INDUCE CONFUSIÓN A LOS MÁS HÁBILES. VER GRÁFICOS PÁG. 397-398 D = p_s - p_i = $\frac{A_{\text{ALTO}} - A_{\text{BAJO}}}{N}$ <table><tr><th>VALORES</th><th>INTERPRETACIÓN</th></tr><tr><td>D ≥ 0,40</td><td>DISCRIMINA MUY BIEN</td></tr><tr><td>0,30 ≥ D ≥ 0,39</td><td>DISCRIMINA BIEN</td></tr><tr><td>0,20 ≥ D ≥ 0,29</td><td>DISCRIMINA POCO</td></tr><tr><td>0,10 ≥ D ≥ 0,19</td><td>NECESITA REVISIÓN</td></tr><tr><td>D < 0,10</td><td>CARECE DE UTILIDAD</td></tr></table>	VALORES	INTERPRETACIÓN	D ≥ 0,40	DISCRIMINA MUY BIEN	0,30 ≥ D ≥ 0,39	DISCRIMINA BIEN	0,20 ≥ D ≥ 0,29	DISCRIMINA POCO	0,10 ≥ D ≥ 0,19	NECESITA REVISIÓN	D < 0,10	CARECE DE UTILIDAD
VALORES	INTERPRETACIÓN												
D ≥ 0,40	DISCRIMINA MUY BIEN												
0,30 ≥ D ≥ 0,39	DISCRIMINA BIEN												
0,20 ≥ D ≥ 0,29	DISCRIMINA POCO												
0,10 ≥ D ≥ 0,19	NECESITA REVISIÓN												
D < 0,10	CARECE DE UTILIDAD												
ÍNDICES DE DISCRIMINACIÓN BASADOS EN LA CORRELACIÓN	A TRAVÉS DE LA CORRELACIÓN ENTRE LAS PUNTUACIONES EN EL ÍTEM Y EN EL TEST TOTAL. SI ES POSITIVA EL ÍTEM DISCRIMINARÁ BIEN. • SE DEBE <u>DESCONTAR DE LA PUNTUACIÓN TOTAL LA PUNTUACIÓN DEL ÍTEM</u>, PARA EVITAR CORRELACIONES ESPÚREAS. • EL ÍNDICE DE CORRELACIÓN HA DE SER COHERENTE CON EL TIPO DE PUNTUACIONES DE ÍTEM Y TEST: COEFICIENTE Φ PUNTUACIONES DICOTÓMICAS ÍTEM Y TEST $\phi = \frac{R_{xy} - p_x p_y}{\sqrt{p_x q_x p_y q_y}}$ CORRELACIÓN BISERIAL-PUNTUAL ÍTEM DICOTÓMICO – TEST CONTINUO $r_{bp} = \frac{\bar{X}_A - \bar{X}_T}{S_x} \sqrt{\frac{p}{q}}$ CORRELACIÓN BISERIAL AMBAS CONTINUAS, PERO UNA SE HA DICOTOMIZADO $r_b = \frac{\bar{X}_A - \bar{X}_T}{S_x} \frac{p}{y}$ RELACIÓN r_{bp} Y r_b $r_{bp} = r_b \frac{y}{\sqrt{pq}}$ COMO $y < \sqrt{pq} \rightarrow$ BISERIAL > BP SE UTILIZA PARA CALCULAR LA DISCRIMINACIÓN EN TEST DE ACTITUDES. AL TRATARSE DE ÍTEMS NO DICOTÓMICOS EL COEFICIENTE ADECUADO SERÍA EL DE PEARSON. SE PUEDE INTERPRETAR COMO UN ÍNDICE DE HOMOGENEIDAD (IH) QUE <u>INDICA HASTA QUÉ PUNTO SE ESTÁN MIDIENDO MISMA DIMENSIÓN ÍTEM Y TEST</u>. APLICAMOS CORRECCIÓN PARA DESCONTAR LA PUNTUACIÓN DEL ÍTEM DEL TOTAL. OTRO PROCEDIMIENTO SERÍA MEDIANTE CONTRASTE DE HIPÓTESIS: QUEREMOS VER <u>SI LA MEDIA EN EL ÍTEM DEL GRUPO DE SUJETOS CON PUNTUACIONES MÁS ALTA EN EL TEST (EL 25% O 27% SUPERIOR) ES SIGNIFICATIVAMENTE SUPERIOR A LA MEDIA EN EL ÍTEM DE SUJETOS CON PUNTUACIONES MÁS BAJAS (25% O 27% INFERIOR)</u> HIPÓTESIS (CONTRASTE UNILATERAL) H₀: $\bar{X}_{sj} = \bar{X}_{ij}$ EL ÍTEM NO DISCRIMINA H₁: $\bar{X}_{sj} \neq \bar{X}_{ij}$ EL ÍTEM DISCRIMINA SI VALOR EMPÍRICO > QUE EL TEÓRICO RECHAZAMOS H₀ $R_{j(x-j)} = \frac{R_{jx} S_x - S_j}{\sqrt{S_x^2 + S_j^2 - 2R_{jx} S_x S_j}}$ ESTADÍSTICO DE CONTRASTE (t STUDENT/MUESTRAS INDEP): $T = \frac{\bar{X}_{sj} - \bar{X}_{ij}}{\sqrt{\frac{(n_s - 1) S_{sj}^2 + (n_i - 1) S_{ij}^2}{n_s + n_i - 2} \left(\frac{1}{n_s} + \frac{1}{n_i} \right)}}$ (n_s + n_i - 2) g.l.												

FACTORES QUE AFECTAN A LA DISCRIMINACIÓN	• VARIABILIDAD. CUANDO UN ÍTEM TIENE VARIANZA CERO (ACERTADO O FALLADO POR TODOS) NO TIENE PODER DE DISCRIMINACIÓN. $S_x = \sum S_j r_{jx}$ PARA ÍTEMS DICOTÓMICOS: $S_x^2 = \sum \sqrt{pq} r_{j(x-j)}$	• DIFICULTAD DEL ÍTEM. MUY IMPORTANTE <u>PARA OPTIMIZAR LA DISCRIMINACIÓN</u>. UN ÍTEM ALCANZA SU MÁXIMO PODER DISCRIMINATIVO CUANDO SU DIFICULTAD ES <u>MEDIA</u> (VARIANZA MÁXIMA CUANDO p = 0,5)	• DIMENSIONALIDAD DEL TEST. EN TEST MULTIDIMENSIONALES EL ÍNDICE DE DISCRIMINACIÓN HAY QUE CALCULARLO RESPECTO AL <u>CONJUNTO DE ÍTEMS ASOCIADOS A UNA MISMA DIMENSIÓN O CONSTRUCTO</u>, YA QUE SI NO LA CORRELACIÓN SE VE AFECTADA A LA BAJA.	• FIABILIDAD DEL TEST. FIABILIDAD Y DISCRIMINACIÓN ESTÁN INTIMAMENTE RELACIONADOS. <u>AL AUMENTAR LA DISCRIMINACIÓN</u> DE LOS ÍTEMS AUMENTA LA FIABILIDAD DEL TEST. USAMOS α DE CRONBACH SUSTITUYENDO S_x POR $\sum S_j r_{jx}$ $\alpha = \frac{n}{n-1} \left[1 - \frac{\sum S_j^2}{(\sum IF)^2} \right]$
---	---	--	---	--

3. FIABILIDAD Y VALIDEZ DEL ÍTEM

ÍNDICE DE FIABILIDAD IF CORRELACIÓN DEL ÍTEM CON UN CRITERIO INTERNO. SE CONSIDERA UNA <u>MEDIDA DE LA PRECISIÓN CON LA QUE EL ÍTEM MIDE EL CONSTRUCTO</u> O DOMINIO DE INTERÉS. • LA FIABILIDAD DEL TEST SE PUEDE EXPRESAR EN TÉRMINOS DE LOS IF. α DE CRONBACH . RELACIÓN DIRECTA ENTRE LA FIABILIDAD DE LOS ÍTEMS Y LA DEL TEST. $IF = S_j r_{jx} = S_j ID_j$	ÍNDICE DE VALIDEZ IV SE EVALÚA LA RELACIÓN DEL ÍTEM CON UN CRITERIO EXTERNO . <u>CORRELACIONAR LAS PUNTUACIONES DE UNA MUESTRA DE SUJETOS EN EL ÍTEM CON LAS PUNTUACIONES OBTENIDAS POR LOS MIMOS SUJETOS CON ALGÚN CRITERIO EXTERNO DE INTERÉS.</u> $IV = S_j r_{jy}$ SI EL ÍTEM ES DICOTÓMICO Y EL CRITERIO CONTINUO: BISERIAL-PUNTUAL: $IV = S_j r_{bpjy}$	EL COEFICIENTE DE VALIDEZ DE UN TEST SE PUEDE EXPRESAR A TRAVÉS DE LOS IV DE LOS ÍTEMS. CUÁNTO MAYORES SEAN LOS IV MÁS OPTIMIZARÁN LA VALIDEZ DEL TEST. $r_{xy} = \frac{\sum IV}{\sum IF}$ PERMITE VER COMO LA VALIDEZ DEL TEST SE PUEDE ESTIMAR A PARTIR DE LA DISCRIMINACIÓN DE CADA UNO DE LOS ÍTEMS, DE SU VALIDEZ Y DE SU VARIABILIDAD. PARADOJA: SI QUEREMOS SELECCIONAR ÍTEMS PARA MAXIMIZAR LA FIABILIDAD DEL TEST, TENDRÍAMOS QUE ELEGIR AQUELLOS CUYO ID SEA ALTO, PERO ESTO NOS LLEVARÍA A REDUCIR EL COEFICIENTE DE VALIDEZ DEL TEST PORQUE ID ESTÁ EN EL DENOMINADOR.
---	---	---

5. ANÁLISIS DE DISTRACTORES

DESPUÉS DE ANALIZAR LA ALTERNATIVA CORRECTA ES IMPORTANTE EL ANÁLISIS DE LAS ALTERNATIVAS INCORRECTAS. ESTE ANÁLISIS IMPLICA <u>INDAGAR EN LA DISTRIBUCIÓN DE LOS SUJETOS A LO LARGO DE LOS DISTRACTORES</u> . • CARACTERÍSTICAS DE LOS BUENOS DISTRACTORES : DEBE SER <u>ELEGIDO POR UN MÍNIMO DE SUJETOS</u> (AL MENOS UN 10%) PERO SI ESTA POR ENCIMA NO ES UN BUEN DISTRACTOR. • DEBEN SER <u>EQUIPROBABLES</u> . PORCENTAJE SIMILAR DE SUJETOS QUE ELIGE CADA DISTRACTOR. • EL RENDIMIENTO EN EL TEST DE LOS SUJETOS QUE HAN SELECCIONADO CADA ALTERNATIVA INCORRECTA DEBE SER MENOR QUE EL DE LOS QUE HAN COGIDO LA CORRECTA. • A MEDIDA QUE AUMENTE EL NIVEL DE APTITUD DE LOS SUJETOS, EL PORCENTAJE DE ELLOS QUE SELECCIONEN LAS ALTERNATIVAS INCORRECTAS DEBE DISMINUIR.	EQUIPROBABILIDAD DE LOS DISTRACTORES PRUEBA DE LA INDEPENDENCIA DE χ^2 HIPÓTESIS: $H_0: FT = FO$ $H_1: FT \neq FO$ $\chi^2 = \sum_{j=1}^k \frac{(FT - FO)^2}{FT}$ FT FRECUENCIAS TEÓRICAS FO FRECUENCIAS OBSERVADAS (k - 1) g.l. k ALTERNATIVAS INCORRECTAS. <u>SI VALOR EMPÍRICO > QUE EL TEÓRICO RECHAZAMOS H_0</u>	PODER DISCRIMINATIVO DE LOS DISTRACTORES UN DISTRACTOR DEBE DISCRIMINAR EN SENTIDO CONTRARIO A LA OPCIÓN CORRECTA: <u>LA CORRELACIÓN DE LA OPCIÓN CORRECTA DEBE SER POSITIVA Y LA DEL DISTRACTOR NEGATIVA</u> . • <u>DISCRIMINACIÓN POSITIVA</u> : CORRELACIÓN POSITIVA ENTRE LA OPCIÓN CORRECTA Y LAS NOTAS DE LOS SUJETOS EN EL TEST. • EN LOS DISTRACTORES SE DA UN <u>DISCRIMINACIÓN NEGATIVA</u> (TENDENCIA CONTRARIA) CUANTO MAYOR NIVEL DE APTITUD DE LOS SUJETOS, LA ELIGEN MENOS. LA MEDIA DEL DISTRACTOR DEBE SER MENOR QUE LA MEDIA DE LA ALTERNATIVA CORRECTA $\bar{X}_{distractor} < \bar{X}_{correcta}$ LA MEDIA DEL DISTRACTOR DEBE SER MENOR QUE LA MEDIA TOTAL $\bar{X}_{distractor} < \bar{X}_{total}$
--	---	--

6. FUNCIONAMIENTO DIFERENCIAL DEL ÍTEM

FDI	SESGO	IMPACTO
<u>ANALIZA SI DE MANERA SISTEMÁTICA DISTINTOS GRUPOS CON EL MISMO NIVEL DE COMPETENCIA TIENEN DISTINTAS PROBABILIDADES DE ÉXITO EN UN ÍTEM, PERO SIN APUNTAR A POSIBLES CAUSAS.</u> • INDICA QUE <u>LAS DIFERENCIAS NO SON REALES</u> . • ES UNA <u>FUENTE DE INVALIDEZ O ERROR SISTEMÁTICO</u> (VALIDEZ DE CONSTRUCTO) CONTRARIO AL ERROR ALEATORIO. • CON EL FDI <u>DETECTAMOS ÍTEMS SESGADOS</u> .	<u>ESTE TÉRMINO SE UTILIZA PARA EL ESTUDIO DE LAS POSIBLES CAUSAS DE DICHO FUNCIONAMIENTO DIFERENCIAL.</u> LAS DIFERENCIAS ENCONTRADAS POR FDI SE PUEDEN DEBER A DIFERENCIAS REALES ENTRE LOS GRUPOS O A FUENTES SISTEMÁTICAS DE VARIACIÓN AJENAS AL CONSTRUCTO. • UN ÍTEM ESTÁ <u>SESGADO</u> CUANDO LA <u>PROBABILIDAD DE ÉXITO</u> PARA SUJETOS IGUALMENTE COMPETENTES EN LA VARIABLE QUE MIDE EL ÍTEM, <u>DEPENDE DEL GRUPO AL QUE PERTENEZCA EL SUJETO</u> . • UN ÍTEM ES <u>INSESGADO</u> SI LA <u>PROBABILIDAD DE ÉXITO</u> PARA EL ÍTEM ES LA MISMA PARA SUJETOS IGUALMENTE CAPACES <u>INDEPENDIENTEMENTE DEL GRUPO AL QUE PERTENEZCA</u> .	UN ÍTEM PRESENTA IMPACTO CUANDO <u>REFLEJA UNA DIFERENCIA REAL (NO UN ERROR SISTEMÁTICO) ENTRE GRUPOS EN LA VARIABLE MEDIDA</u> . ES LÍCITO QUE DEBIDO A DIFERENCIAS EN EL NIVEL DE COMPETENCIA DE SUBPOBLACIONES, LA PROBABILIDAD DE RESPONDERLO CORRECTAMENTE SEA MAYOR PARA UN GRUPO QUE PARA OTRO.

MANTEL - HAENSZEL

<u>PROCEDIMIENTO:</u>		<u>HIPÓTESIS:</u> UN ÍTEM NO PRESENTARÁ FDI SI EL COCIENTE ENTRE LOS SUJETOS QUE ACIERTAN EL ÍTEM Y LOS QUE LO FALLAN ES EL MISMO PARA LOS DOS GRUPOS EN CADA UNO DE LOS NIVELES DE APTITUD, ES DECIR:																		
1.	IDENTIFICAR UNA <u>VARIABLE</u> QUE SEA LA POSIBLE <u>CAUSA</u> .	$H_0: \frac{A_i}{B_i} = \frac{C_i}{D_i}$																		
2.	<u>DOS GRUPOS:</u> <u>REFERENCIA</u> (GR; NORMALMENTE EL FAVORECIDO) Y <u>FOCAL</u> (GF; EL PERJUDICADO)																			
3.	SE ESTABLECEN <u>DISTINTOS NIVELES DE APTITUD</u> TOMANDO LA PUNTUACIÓN EMPÍRICA OBTENIDA EN EL TEST.																			
4.	<u>SE CUENTA EL NÚMERO DE RESPUESTAS CORRECTAS E INCORRECTAS POR CADA GRUPO Y NIVEL DE HABILIDAD i.</u>	<u>ESTADÍSTICO DE MANTEL-HAENSZEL:</u>	<table><tr><th></th><th>Correctas</th><th>Incorrectas</th><th></th></tr><tr><th>GR</th><td>A_i</td><td>B_i</td><td>n_{RI}</td></tr><tr><th>GF</th><td>C_i</td><td>D_i</td><td>n_{FI}</td></tr><tr><th></th><td>n_{1i}</td><td>n_{2i}</td><td>N_i</td></tr></table>			Correctas	Incorrectas		GR	A _i	B _i	n _{RI}	GF	C _i	D _i	n _{FI}		n _{1i}	n _{2i}	N _i
	Correctas	Incorrectas																		
GR	A _i	B _i			n _{RI}															
GF	C _i	D _i	n _{FI}																	
	n _{1i}	n _{2i}	N _i																	
		$\alpha_{MH} = \frac{\sum_{i=1}^n \frac{A_i D_i}{N_i}}{\sum_{i=1}^n \frac{B_i C_i}{N_i}}$																		

7. CONDICIONES NECESARIAS QUE DEBE SATISFACER UN ÍTEM

DIFICULTAD	• <u>EN TEST DE EJECUCIÓN MÁXIMA</u>: <u>0,20 < ID < 0,80</u>. LA MAYORÍA DE ELLOS DEBERÍAN PRESENTAR NIVELES MEDIOS DE DIFICULTAD: <u>0,30 < ID < 0,70</u>. • <u>EN ÍTEMS DE ACTITUDES</u>: NO ES TAN IMPORTANTE ESTE PARÁMETRO. SE TRADUCE EN EL <u>GRADO DE ACTITUD MEDIA ANTE EL ÍTEM</u>. DEPENDE DE SI ES UNA ACTITUD POSITIVA O NEGATIVA SERÁ BUENO O MALO OBTENER UN VALOR MEDIO ALTO.	DISTRACTORES	<u>DEBEN FUNCIONAR COMO TALES</u> . <u>CADA ALTERNATIVA INCORRECTA DEBE SER SELECCIONADA POR BASTANTES MÁS SUJETOS CON PUNTUACIONES BAJAS EN EL TEST</u> QUE POR AQUELLOS OTROS QUE PRESENTAN UN BUEN NIVEL DE APTITUD. Y ADEMÁS <u>DEBEN SER EQUIPROBALES</u> .
DISCRIMINACIÓN	• <u>UN ÍTEM DEBE DISCRIMINAR CLARAMENTE ENTRE GRUPOS ALTOS Y BAJOS EN APTITUD Y ACTITUD</u>. CUÁNTO MÁS DISCRIMINE MEJOR. POR ENCIMA DE <u>0,30 EN LOS DE APTITUDES</u>. Y DE <u>0,20 EN LOS DE ACTITUDES</u>.	FDI	CUANDO LOS SUJETOS TIENEN EL <u>MISMO NIVEL DE RASGO Y PRESENTAN DISTINTA PROBABILIDAD DE ACERTAR</u> UN DETERMINADO ÍTEM, ES NECESARIO LLEVAR A CABO UN <u>ANÁLISIS EXHAUSTIVO</u> POR SI PRESENTARA UN FUNCIONAMIENTO DIFERENCIAL.

COMPARACIÓN DE LAS RESPUESTAS DE LOS ÍTEM

PARA MEDIR EL GRADO DE HOMOGENEIDAD DE DOS ÍTEMS PODEMOS APLICAR EL ESTADÍSTICO χ^2 CON 1 g.l. $\chi^2 = \frac{n(ad - cd)^2}{(a + b)(c + d)(b + d)(a + c)}$ • H_0: <u>EL GRADO DE ACUERDO U HOMOGENEIDAD ENTRE LOS ÍTEMS NO ES SIGNIFICATIVO</u> (LOS DOS ÍTEMS NO MIDEN LA MISMA VARIABLE, UNO DE ELLOS O LOS DOS SON INADECUADOS)	PODEMOS PLANTEAR ADEMÁS SI ES GRADO DE DIFICULTAD DE ESTOS DOS ÍTEMS ES IGUAL O NO, ES DECIR SI LAS DIFERENCIAS ENCONTRADAS SON LO SUFICIENTEMENTE PEQUEÑAS COMO PARA SER DEBIDAS A ERRORES DE MUESTREO <u>H_0: LA EJECUCIÓN DE AMBOS ÍTEMS ES IGUAL</u> $\chi^2 = \frac{(b - c - 1)^2}{b + c}$									
• CUANDO LA HOMOGENEIDAD ES SIGNIFICATIVA, SE CALCULA EL GRADO O INTENSIDAD DE DICHA HOMOGENEIDAD COMO LA PROPORCIÓN DE SUJETOS QUE CONTESTARON CORRECTAMENTE A AMBOS ÍTEMS (ESTE ÍNDICE REPRESENTA LA CONSISTENCIA ENTRE LOS ACIERTOS)	$p = \frac{(a + d)}{n}$ <table><tr><td></td><td>ACIERTOS</td><td>ERRORES</td></tr><tr><td>ACIERTOS</td><td>A</td><td>B</td></tr><tr><td>ERRORES</td><td>C</td><td>D</td></tr></table>		ACIERTOS	ERRORES	ACIERTOS	A	B	ERRORES	C	D
	ACIERTOS	ERRORES								
ACIERTOS	A	B								
ERRORES	C	D								

TEMA 9. ASIGNACIÓN, TRANSFORMACIÓN Y EQUIPARACIÓN DE LAS PUNTUACIONES

NECESIDAD DE TRANSFORMACIÓN DE LAS PUNTUACIONES PARA SU INTERPRETACIÓN

CONSTRUCCIÓN DE TEST COMO INSTRUMENTO CIENTÍFICO DE MEDICIÓN. ETAPAS

1. INICIAL: **CONSTRUCCIÓN (TEMAS 2 Y 3)**
2. SEGUNDA: **EVALUACIÓN (TEMAS 8 Y 4, 5, 6 Y 7)**
3. TERCERA: **APLICACIÓN Y EVALUACIÓN A LOS SUJETOS (TEMA 9)**
4. FINAL: **MANUAL DEL TEST (TEMA 9)**

ESTAMOS EN LA FASE 3 DE LA CONSTRUCCIÓN DE UN TEST. UNA VEZ ELABORADA LA PRUEBA DEFINITIVA HAY QUE **APLICARLA, ASIGNAR PUNTUACIONES A LOS SUJETOS Y DOTAR DE SIGNIFICADO A ESAS PUNTUACIONES PARA PODERLAS INTERPRETAR.**

LAS PUNTACIONES EMPÍRICAS SE TRANSFORMAN PARA CONSEGUIR INFORMACIÓN MAS FÁCILMENTE COMPRENSIBLE TANTO PARA LOS SUJETOS QUE HAN REALIZADO EL TEST COMO PARA TODO AQUEL QUE QUIERA CONOCER SU SIGNIFICADO.

LA PUNTUACIÓN OBTENIDA POR UN SUJETO ES UNA **DESCRIPCIÓN CUANTITATIVA DEL RASGO** QUE ESTAMOS EVALUANDO.

INTERPRETACIÓN NORMATIVA: REFERIDA A LA NORMA SE COMPARA LA PUNTUACIÓN OBTENIDA POR UN SUJETO EN UN TEST CON LAS OBTENIDAS, EN EL MISMO TEST, POR UN **GRUPO DE REFERENCIA O GRUPO NORMATIVO.**

- **NORMAS:** PUNTUACIONES OBTENIDAS POR LOS SUJETOS Y SUS TRANSFORMACIONES.
- **BAREMO DEL TEST:** CONJUNTO DE TODAS LAS NORMAS.

INTERPRETACIÓN CRITERIAL: REFERIDA AL CRITERIO EL INTERÉS CENTRAL ES DETERMINAR EL GRADO DE DOMINIO QUE UN SUJETO TIENE SOBRE UN **CRITERIO PREESTABLECIDO.** SE SUELE TOMAR UNA **PUNTUACIÓN DE CORTE** QUE PERMITA CLASIFICAR A LAS SUJETOS EN DOS GRUPOS. Y ANALIZAR LOS POSIBLES ERRORES DE CLASIFICACIÓN.

1. TRANSFORMACIÓN DE LAS PUNTUACIONES EN LOS TEST REFERIDOS A NORMAS

DADO QUE SE TRATA DE UNA INTERPRETACIÓN NORMATIVA, ES NECESARIO SELECCIONAR DE LA POBLACIÓN OBJETO DE ESTUDIO UNA **MUESTRA REPRESENTATIVA** A LA QUE SE APLICA EL/LOS TEST. SOBRE ESA MUESTRA SE OBTIENEN TODAS LAS **NORMAS.** UNA VEZ ESTABLECIDAS SE PUEDE COMPARAR LA PUNTUACIÓN OBTENIDA POR UN SUJETO DE LA MISMA POBLACIÓN PARA SABER CUÁL ES SU **POSICIÓN** RESPECTO A LA DEL GRUPO NORMATIVO

TRANSFORMACIONES		LINEALES
		TÍPICAS
	LINEALES	TÍPICAS NOS INDICA EL NÚMERO DE DESVIACIONES TÍPICAS A LAS QUE SE ENCUENTRA LA PUNTUACIÓN DE UN SUJETO RESPECTO DE LA MEDIA DEL GRUPO NORMATIVO. $Z_x = \frac{X - \bar{X}}{S_x}$ INCONVENIENTE. TIENE DE MEDIA 0 Y DESVIACIÓN TÍPICA 1. LA DISTRIBUCIÓN DE PUNTUACIONES TÍPICAS DE UNA VARIABLE NORMAL SUELE OSCILAR DE -3 A +3, LO QUE IMPLICA VALORES NEGATIVOS Y DECIMALES. UNA FORMA DE SOLUCIONAR ESTE INCONVENIENTE SON LAS ESCALAS TÍPICAS DERIVADAS.
	NO LINEALES	
		TÍPICAS DERIVADAS
		$Y = a + bZ_x$ ESCALA D: $D = 50 + 20Z_x$ MEDIA IGUAL A 50, DESVIACIÓN TÍPICA 20. ESCALA T: $T = 50 + 10Z_x$ MEDIA IGUAL A 50, DESVIACIÓN TÍPICA 10.
		INCONVENIENTE. PERSISTE UN PROBLEMA TAMBIÉN COMÚN A LA ESCALA TÍPICA: LA APLICACIÓN DE UN TEST A DISTINTAS MUESTRAS DE SUJETOS DARÁ LUGAR SEGURAMENTE, A VALORES DISTINTOS TANTO DE LA MEDIA COMO LA DESVIACIÓN TÍPICA Y, EN ALGUNOS CASOS, LAS DISTRIBUCIONES NO SERÁN IGUALES. CUIDADO AL COMPARAR UN SUJETO CON UNA DETERMINADA MUESTRA. PARA CAMBIAR LA FORMA DE LA DISTRIBUCIÓN: TÍPICAS NORMALIZADAS.

NO LINEALES		CRONOLÓGICAS
PERCENTILES	TÍPICAS NORMALIZADAS	LA INTERPRETACIÓN DE LA PUNTUACIÓN OBTENIDA POR UN SUJETO EN UN TEST, SE LLEVA A CABO CON <u>RELACIÓN A SU EDAD Y CON LA PUNTUACIÓN MEDIA OBTENIDA POR LOS SUJETOS DE SU EDAD</u> . <u>INCONVENIENTE</u> . NO SON MUY RECOMENDABLES. <u>NO SIEMPRE ES POSIBLE LA COMPARACIÓN</u> . EL <u>SIGNIFICADO DE UN AÑO DE EDAD MENTAL NO ES CONSTANTE</u> CON EL DESARROLLO EVOLUTIVO DEL NIÑO.
AQUELLA <u>PUNTUACIÓN DEL TEST QUE DEJA POR DEBAJO DE SÍ UN DETERMINADO PORCENTAJE DE CASOS DEL GRUPO NORMATIVO</u> . NOS PROPORCIONA UNA IDEA DE <u>POSICIÓN</u> . CONSTITUYEN UNA ESCALA ORDINAL.	SE OBTIENEN A PARTIR DE LOS PERCENTILES. MEDIANTE LA TABLA DE LA CURVA NORMAL, SE BUSCA EL VALOR DE LA PUNTUACIÓN TÍPICA Z_n . ES LA <u>PUNTUACIÓN TÍPICA QUE LE CORRESPONDE A UNA PUNTUACIÓN EMPÍRICA</u> OBTENIDA POR UN SUJETO EN UN TEST EN UNA <u>DISTRIBUCIÓN NORMAL</u> .	EDAD MENTAL
$P_x o C_x = \frac{100}{N} \left(f_b + \frac{f_d}{I} (X_c - L_i) \right) = f_{ac} \frac{100}{N}$	NORMALIZADAS DERIVADAS	PROPUESTAS EN PRINCIPIO POR BINET Y LUEGO POR BINET- SIMON. SELECCIÓN DE MUESTRAS DE NIÑOS DE CADA RANGO DE EDAD A LOS QUE SE LES APLICA EL TEST PARA CALCULAR PUNTUACIÓN MEDIA. TABLA CON LAS CORRESPONDENCIAS. <u>ASIGNAMOS EDAD MENTAL INDEPENDIENTE/ DE LA EDAD CRONOLÓGICA</u> .
PORCENTAJE QUE LE CORRESPONDE A UN SUJETO QUE SUPERA AL X% DE LOS SUJETOS:	ESTANINOS O ENEATIPOS . <u>ESCALA DE VALORES ENTEROS Y POSITIVOS DE 9 UNIDADES</u> , DEL 1 AL 9. $E = 5 + 2(Z_n)$	COCIENTE INTELECTUAL
$X_c = L_i + \left(\frac{N \cdot C_x}{100} - f_b \right) \frac{I}{f_d}$	CADA ESTANINO ACUMULA UN PORCENTAJE DE DISTRIBUCIÓN NORMAL Y DE PERCENTILES. TABLA DE EQUIVALENCIA. <u>INCONVENIENTE</u> : AL INCLUIR EN EL MISMO ENEATIPO A SUJETOS CON DISTINTAS PUNTUACIONES, <u>SE PIERDE BASTANTE INFORMACIÓN</u> .	$CI = \frac{EM}{EC} \cdot 100$ <u>INCONVENIENTE</u> : <u>POCO DISCRIMINATIVO EN ADULTOS</u> DEBIDO A QUE LA EM SE ESTABILIZA. LAS DISTRIBUCIONES DE CI PARA DISTINTAS EDADES TIENEN DISTINTAS DESVIACIONES TÍPICAS. ESTO IMPLICA QUE <u>EL MISMO CI NO PROPORCIONA LA MISMA POSICIÓN RELATIVA</u> EN LAS DISTRIBUCIONES DE DISTINTAS EDADES.
ES UNA DE LAS PUNTUACIONES DE MAYOR USO POR SU FACILIDAD DE INTERPRETACIÓN. SU SIGNIFICADO ES EL MISMO INDEPENDIENTEMENTE DEL TEST APLICADO Y DE LA FORMA DE LA DISTRIBUCIÓN. NOS PERMITEN <u>COMPARAR</u> PUNTUACIONES DE SUJETOS DISTINTOS EN UN MISMO TEST <u>SIEMPRE QUE PERTENEZCAN AL MISMO GRUPO DE EDAD (INCONVENIENTE)</u> .		$SI \ EM = \ EC \Rightarrow CI = 100$ $SI \ EM < \ EC \Rightarrow CI < 100$ $SI \ EM > \ EC \Rightarrow CI > 100$

2. EQUIPARACIÓN DE PUNTUACIONES

PROCESO MEDIANTE EL CUAL SE ESTABLECE UNA CORRESPONDENCIA ENTRE LAS PUNTUACIÓN DE VARIOS TEST, DE TAL MANERA QUE SEA INDISTINTO EL EMPLEO DE UNO U OTRO, PUESTO QUE LAS PUNTUACIONES DE CUALQUIERA DE ELLOS SE PODRÁN EXPRESAR EN TÉRMINOS DE LAS DEL OTRO TEST.

2 CONDICIONES FUNDAMENTALES:

- * LOS TESTS DEBEN MEDIR EL MISMO CONSTRUCTO PSICOLÓGICO.
- * LOS TESTS DEBEN TENER LA MISMA FIABILIDAD.

2 TIPOS DE EQUIPARACIÓN:

- * EQUIPARACIÓN HORIZONTAL: LOS DOS TESTS TIENEN LA MISMA DIFICULTAD.
- * EQUIPARACIÓN VERTICAL: LOS DOS TESTS TIENEN DIFERENTE DIFICULTAD.

PARA HACER UN ESTUDIO DE EQUIPARACIÓN ES IMPORTANTE QUE EL NÚMERO DE SUJETOS QUE SE VAYA A UTILIZAR SEA REPRESENTATIVO DE LA POBLACIÓN A LA QUE VA A IR DESTINADO EL TEST.

DISEÑOS DE EQUIPARACIÓN

DE UN SOLO GRUPO	DE GRUPOS EQUIVALENTES	DE GRUPOS NO EQUIVALENTES CON ÍTEMS COMUNES. DISEÑO ANCLAJE
SE ADMINISTRAN <u>DOS FORMAS DEL TEST</u> AL <u>MISMO GRUPO DE SUJETOS</u> . DEBEN MEDIR MISMO RASGO Y TENER MISMO NIVEL DE DIFICULTAD. SE APLICA CONTRABALANCEADO PARA EVITAR POSIBLES EFECTOS DE ORDEN DE APLICACIÓN: <u>DISEÑO DE UN SOLO GRUPO CONTRABALANCEADO</u> . LAS DOS APLICACIONES DEBEN SER INDEPENDIENTES.	SE FORMAN <u>DOS GRUPOS AL AZAR</u> Y SE APLICA <u>A CADA UNO DE ELLOS UNA DE LAS FORMAS</u> DEL TEST. EVITA LOS EFECTOS DE LA FATIGA, APRENDIZAJE U ORDEN DE APLICACIÓN. IMPORTANTE QUE AMBOS <u>GRUPOS SEAN EQUIVALENTES EN LA APTITUD</u> QUE MIDE EL TEST PARA <u>EVITAR SESGOS EN EL PROCESO DE EQUIPARACIÓN</u> .	ES EL DISEÑO MÁS UTILIZADO. ES IGUAL QUE EL ANTERIOAR PERO EN ESTE CASO <u>LAS MUESTRAS NO DEBEN SER EQUIVALENTES</u> Y ADEMÁS A LAS DOS MUESTRAS SE LES APLICA <u>UN TEST COMÚN (Z): TEST DE ANCLAJE QUE PERMITE ESTABLECER LAS EQUIVALENCIAS</u> . DOS MODALIDADES: <u>ANCLAJE INTERNO</u> . ÍTEMS COMUNES A AMBOS TEST Y ESTOS APARECEN <u>INTERCALADOS</u> CON EL RESTO DE LOS ÍTEMS DE LAS DOS FORMAS X E Y (<u>ÍTEMS DE ANCLAJE</u>) <u>ANCLAJE EXTERNO</u> : LOS ÍTEMS COMUNES APARECEN FORMANDO UN <u>TEST INDEPENDIENTE (TEST DE ANCLAJE)</u> SUS PUNTUACIONES NO UTILIZAN EN EL CÓMPUTO DE LA PUNTUACIÓN TOTAL DE LO SUJETOS EN LAS FORMAS A EQUIPARAR. NÚMERO DE ÍTEMS COMUNES : AL MENOS EL <u>20% DE LA LONGITUD</u> TOTAL DE UN TEST COMPUESTO POR <u>40 ÍTEMS. EN TEST LARGOS CON 30 ES SUFICIENTE</u> . LAS DIFERENCIAS SE PUEDEN DEBER A <u>DIFERENCIAS ENTRE AMBOS GRUPOS</u> DE SUJETOS, O BIEN A LAS DIFERENCIAS <u>ENTRE AMBAS FORMAS</u> .

MÉTODOS DE EQUIPARACIÓN

CON PERCENTILES	DE LA MEDIA	LINEAL
ES EL <u>MÉTODO MÁS HABITUAL</u>. OBTENIDOS LOS PERCENTILES CORRESPONDIENTES A LAS PUNTUACIONES PARA CADA UNO DE LOS TEST, <u>CONSISTE EN EQUIPARAR LAS PUNTUACIONES CUYOS PERCENTILES SON IGUALES</u>. POSTERIORMENTE <u>REPRESENTAMOS GRÁFICAMENTE</u> LAS DOS DISTRIBUCIONES DE PERCENTILES Y DE ESE GRÁFICO OBTENEMOS LAS PUNTUACIONES EQUIVALENTES. PAG.472. IGUAL PARA EL DISEÑO DE UN SOLO GRUPO Y GRUPOS EQUIVALENTES.	<u>HACER CORRESPONDER LAS MEDIAS DE LOS DOS TEST A EQUIPARAR</u>. SE ASUME QUE LAS PUNTUACIONES OBTENIDAS POR UNA MUESTRA DE SUJETOS EN UNO DE LOS TEST DIFIEREN EN UNA <u>CUANTÍA CONSTANTE</u> DE LAS PUNTUACIONES OBTENIDAS POR UNA MUESTRA DE SUJETOS EN EL OTRO TEST. <u>CONSISTE EN SUMAR A UNA PUNTUACIÓN LA DIFERENCIA ENTRE LAS MEDIAS DE LOS DOS TESTS</u>. SEAN X E Y DOS TEST DISTINTOS: $X^* = Y = X - \bar{X} + \bar{Y}$ X* MEDIA DEL TEST Y EQUIVALENTE A UNA DEL TEST X. X PUNTUACIÓN DEL TEST X. $\bar{X}, \bar{Y}$ MEDIAS DE LOS TEST. ES EL MENOS UTILIZADO.	AL CONTRARIO QUE EN EL DE LA MEDIA, EN ESTE CASO SE CONSIDERA QUE <u>LAS DIFERENCIAS ENTRE LAS PUNTUACIONES PUEDEN VARIAR</u>. <u>SE EQUIPARAN AQUELLAS PUNTUACIONES DIRECTAS X E Y CUYAS PUNTUACIONES TÍPICAS SON IGUALES</u> $Z_x = Z_y$ LA FORMA DE CALCULAR LAS MEDIAS Y DESVIACIONES TÍPICAS DE CADA TEST DEPENDE DEL TIPO DE DISEÑO QUE ESTEMOS APLICANDO EN LA RECOGIDA DE DATOS. VER FORMULARIO. DISEÑO DE GRUPOS EQUIVALENTES: $\frac{X - \bar{X}}{S_x} = \frac{Y - \bar{Y}}{S_y} \Rightarrow X^* = Y = \frac{S_y}{S_x} (X - \bar{X}) + \bar{Y}$ DISEÑO DE UN SOLO GRUPO $X^* = Y = \frac{\sqrt{\frac{S_{y1}^2 + S_{y2}^2}{S_{x1}^2 + S_{x2}^2}}}{\frac{\bar{Y}_1 + \bar{Y}_2}{2}} \left(X - \frac{\bar{X}_1 + \bar{X}_2}{2} \right)$ DISEÑO DE ANCLAJE $X^* = Y = \frac{\sqrt{S_{y2}^2 + b_{xz2}(S_z^2 + S_{z2}^2)}}{\sqrt{S_{y1}^2 + b_{xz1}(S_z^2 + S_{z1}^2)} + (\bar{Y}_2 + b_{xz2}(\bar{Z} - \bar{Z}_2))} [X - (\bar{X}_1 + b_{xz1}(\bar{Z} - \bar{Z}_1))]$

4. EL MANUAL DEL TEST

EL MANUAL ES LA INFORMACIÓN QUE SE PRESENTA A LOS USUARIOS DEL MISMO, DE MODO QUE PUEDAN DAR UNA SIGNIFICACIÓN ADECUADA A LA PUNTUACIÓN OBTENIDA POR UN SUJETO EN EL TEST. DETERMINA PUES LA <u>ÓPTIMA COMPRENSIÓN Y APLICACIÓN</u> DEL TEST. SEGÚN YELA EN EL MANUAL DEBEN FIGURAR TODOS LOS DATOS QUE HACEN DEL TEST UN INSTRUMENTO CIENTÍFICO. PODEMOS AGRUPARLOS EN 4 CATEGORÍAS: <u>LA ESPECIFICACIÓN DEL TEST</u> DENOMINACIÓN Y CLASIFICACIÓN <i>EN FUNCIÓN DEL CONSTRUCTO PSICOLÓGICO</i>, EN LA FORMA EN QUE SE PRESENTA EL MATERIAL QUE EMPLEAMOS EN EL TEST (IMPRESO O MANIPULATIVO) O SEGÚN EL MÉTODO DE ADMINISTRACIÓN DE LA PRUEBA (INDIVIDUAL Y COLECTIVA) <u>LA DESCRIPCIÓN DEL TEST</u> INFORMACIÓN RELATIVA A LOS FINES Y FORMA DE APLICACIÓN DEL TEST. 1. <i>OBJETIVO DEL TEST Y CARACTERÍSTICAS</i> 2. <i>CAMPO DE APLICACIÓN</i> (ASPECTOS PSICOLÓGICOS QUE SE PRETENDEN ESTUDIAR, ÁREAS DE APLICACIÓN Y OTRAS APLICACIONES QUE SE HAYAN HECHO CON ÉL Y LOS RESULTADOS OBTENIDOS) 3. <i>DESCRIPCIÓN DETALLADA DEL MATERIAL</i>, QUE INCLUYE: MATERIAL BÁSICO Y MATERIAL AUXILIAR.		1. <i>INSTRUCCIONES DE APLICACIÓN</i> (VITAL IMPORTANCIA) PAUTAS DE CÁRACTER GENERAL SOBRE ATENCIÓN, INTERÉS, COMPRENSIÓN... INSTRUCCIONES ESPECÍFICAS SOBRE SU EJECUCIÓN. TIEMPOS EXACTOS. 2. <i>FORMA DE PUNTUAR</i>. PLANTILLAS CON SOLUCIONES. EN CASO DE TENER PRUEBAS DE CÁRACTER MANIPULATIVO LA MANERA EXACTA DE CONSIDERAR UNA RESPUESTA COMO ACIERTO O ERROR, Y LA FORMA DE CRONOMETRAR. <u>LA JUSTIFICACIÓN DEL TEST</u>. LOS DATOS CUANTITATIVOS Y EXPERIMENTALES QUE JUSTIFICAN EL USO DEL TEST. INFORMACIÓN RELATIVA A LA DURACIÓN DE LA PRUEBA, LA FIABILIDAD, VALIDEZ Y TIPIFICACIÓN DEL TEST. <u>REFERENCIAS BIBLIOGRÁFICAS</u>. INCLUYEN CUALQUIER INFORMACIÓN REFERIDA AL TEST. VER LIBRO: <i>NORMAS PROPUESTAS POR LA APA</i>.
--	--	---