



TEORÍA CLÁSICA DE MEDICIÓN TC Y TEORÍA DE RESPUESTA AL ITEM TRI



UNIVERSIDAD DE CHILE
VICERRECTORÍA DE ASUNTOS
ACADÉMICOS
Departamento de Evaluación, Medición y
Registro Educacional DEMRE

ABRIL DE 2005



TEORÍAS DE LA MEDICIÓN

- SIRVEN COMO MARCO TEÓRICO PARA EL DISEÑO E IMPLEMENTACIÓN DE INSTRUMENTOS DE MEDICIÓN.
- PROPORCIONAN MÉTODOS O PROCEDIMIENTOS PARA DETERMINAR LAS CARACTERÍSTICAS DE LOS ESTÍMULOS O PREGUNTAS QUE FORMAN LAS PRUEBAS.
- A PARTIR DE LAS CARACTERÍSTICAS OBTENIDAS SE DERIVAN MÉTODOS PARA EFECTUAR OTROS ANÁLISIS QUE SEAN DE INTERÉS PARA LOS USUARIOS DE LAS PRUEBAS.

PRINCIPALES TEORÍAS DE LA MEDICIÓN USADAS EN EDUCACIÓN

- TEORÍA CLÁSICA DE LA MEDICIÓN TC.
- TEORÍA DE RESPUESTA AL ÍTEM IRT O TRI.



PARÁMETROS O DESCRIPTORES QUE CARACTERIZAN A UNA PREGUNTA O ÍTEM DE UNA PRUEBA

- **DIFICULTAD:** CUAN FÁCIL O DIFÍCIL RESULTÓ LA PREGUNTA PARA LA POBLACIÓN A LA CUAL SE LE APLICÓ.
- **DISCRIMINACIÓN :** CAPACIDAD QUE TIENE LA PREGUNTA PARA DISTINGUIR, PARA SEPARAR, A LAS PERSONAS QUE TIENEN LA CAPACIDAD QUE SE MIDE DE LAS QUE NO LA POSEEN.

HABILIDAD: EN MEDICIÓN EDUCACIONAL SUELE DENOMINARSE ASI AL RASGO QUE SE DESEA MEDIR CON LA PRUEBA, ESTE PUEDE SER COMPRENSIÓN LECTORA, RAZONAMIENTO CIENTÍFICO, OPERATORIA CON NÚMEROS REALES, ETC.





ASPECTOS COMUNES DE LA TEORÍA CLÁSICA TC Y TEORÍA DE RESPUESTA AL ITEM TRI

AMBAS TEORÍAS PARTEN DE LOS SIGUIENTES PRINCIPIOS :

1. QUE SE HAN DEFINIDO RIGUROSAMENTE:

- ❖ LA VARIABLE QUE SE QUIERE MEDIR.**
- ❖ LOS INDICADORES QUE SERVIRÁN DE BASE PARA LA CONSTRUCCIÓN DE LOS ESTÍMULOS.**

2. QUE LA CONSTRUCCIÓN DE LOS ESTÍMULOS HA SEGUIDO TODAS LAS NORMAS RECOMENDADAS.

3. QUE LA APLICACIÓN EXPERIMENTAL HA SIDO DISEÑADA CONSIDERANDO TODAS LAS CARACTERÍSTICAS DE LA POBLACIÓN OBJETIVO Y LA MUESTRA TIENE UN TAMAÑO ADECUADO.



TEORÍA CLÁSICA

- **INDICADOR DE LA HABILIDAD:** ES EL PUNTAJE QUE SE OBTIENE EN LA PRUEBA, CONSTRUIDO A PARTIR DEL NÚMERO DE RESPUESTAS CORRECTAS O DE ALGUNA FORMA PREVIAMENTE DEFINIDA
- **DIFICULTAD:** PROPORCIÓN DE PERSONAS QUE CONTESTA CORRECTAMENTE UNA PREGUNTA.
- **DISCRIMINACIÓN:** CORRELACIÓN ENTRE CONTESTAR CORRECTA O INCORRECTAMENTE AL ÍTEM Y EL PUNTAJE TOTAL EN LA PRUEBA.

SUPUESTOS DE LA TC

LA TC SUPONE QUE EL PUNTAJE QUE OBTIENE UNA PERSONA EN UNA PRUEBA TIENE DOS COMPONENTES INDEPENDIENTES ENTRE SÍ Y QUE SE SUMAN:

- EL PUNTAJE VERDADERO.
- EL ERROR DE MEDICIÓN.

EN CONSECUENCIA, EL PUNTAJE OBTENIDO ES UNA ESTIMACIÓN DEL VALOR VERDADERO DEL RASGO QUE SE MIDE.





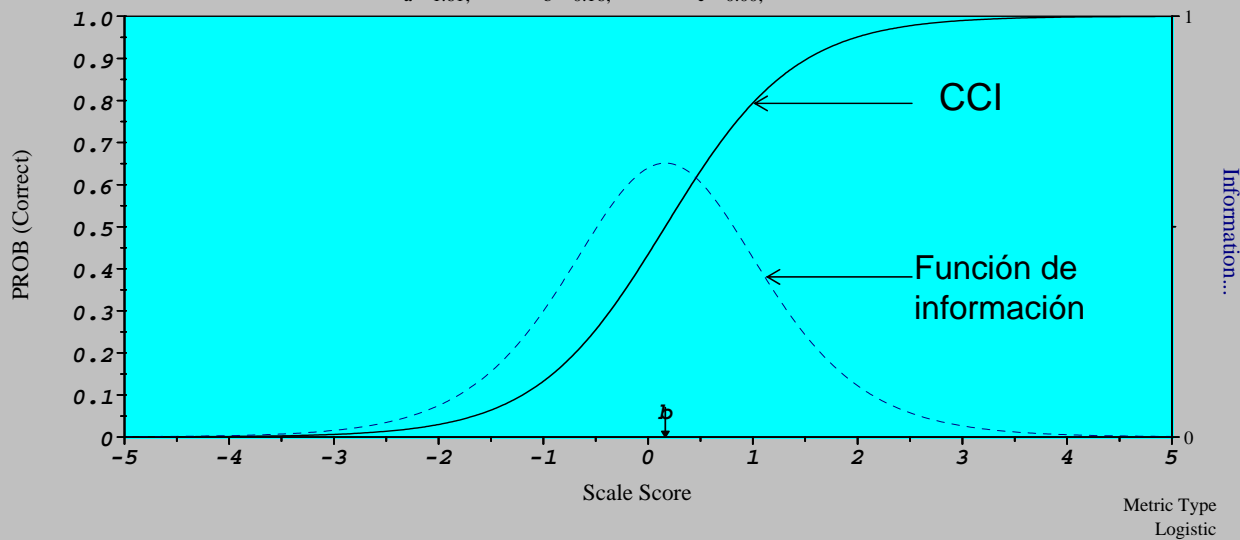
TEORÍA DE RESPUESTA AL ÍTEM TRI

- **LOS MODELOS TRI CENTRAN SU INTERÉS EN LAS PREGUNTAS.**
- **INTENTAN ESTABLECER PARA CADA ÍTEM LA PROBABILIDAD DE SER CONTESTADO CORRECTAMENTE. CURVA CARACTERÍSTICA DE UN ÍTEM CCI.**
- **LA PROBABILIDAD DE UNA RESPUESTA CORRECTA DEPENDE DE LA HABILIDAD DEL EXAMINANDO, Y DE CARACTERÍSTICAS PROPIAS DE LAS PREGUNTAS, TALES COMO: DIFICULTAD, DISCRIMINACIÓN Y PROBABILIDAD DE SER ACERTADA POR AZAR POR UN SUJETO DE MUY BAJA HABILIDAD.**
- **PROPORCIONA INFORMACIÓN SOBRE EL NIVEL DE PRECISIÓN QUE UNA PREGUNTA APORTA EN LA ESTIMACIÓN DE UN DETERMINADO NIVEL DE HABILIDAD. ESTO SE CONOCE CON EL NOMBRE DE “INFORMACIÓN DE UN ÍTEM”**
- **MIENTRAS MAYOR ES LA INFORMACIÓN QUE PROPORCIONA UNA PREGUNTA EN UN DETERMINADO NIVEL DE HABILIDAD, MEJOR ES EL GRADO DE PRECISIÓN CON QUE SE ESTIMA ESE NIVEL DE HABILIDAD. ESTO PERMITE CONSTRUIR PRUEBAS MUY AJUSTADAS AL PROPÓSITO QUE SE PERSIGUE.**



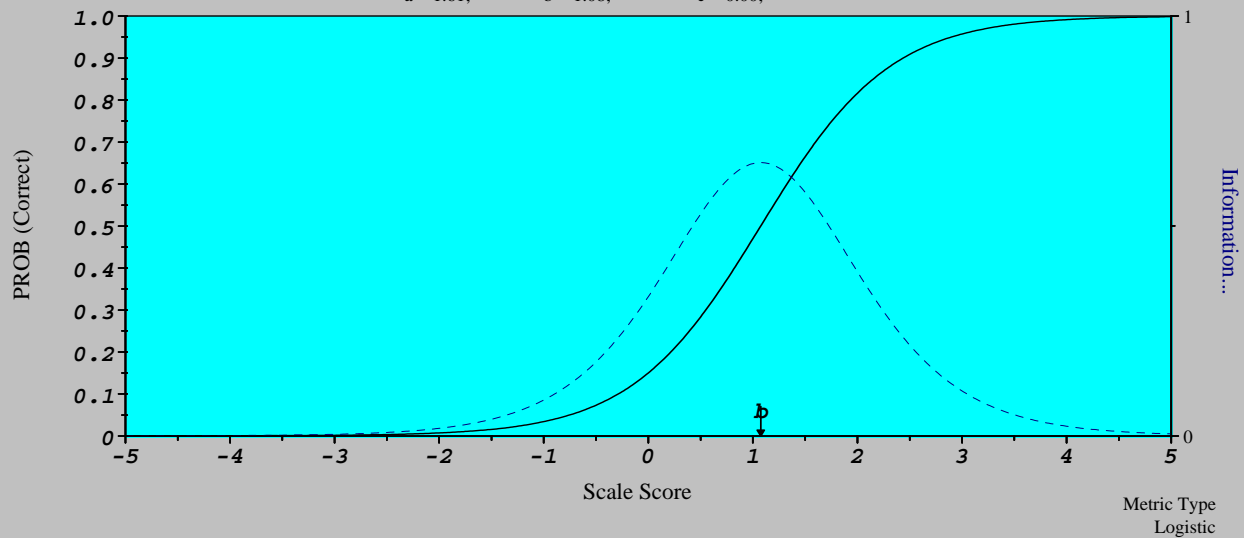
Item Response Function and Item Information

Subtest 1: PAAM10 ; Item 15: 0015
a = 1.61; b = 0.16; c = 0.00;



Item Response Function and Item Information

Subtest 1: PAAM10 ; Item 45: 0045
a = 1.61; b = 1.08; c = 0.00;





SUPUESTOS CLAVES DE LA TRI

- ❖ **UNIDIMENSIONALIDAD** : LA PUNTUACIÓN DE UNA PERSONA EN EL TEST DEPENDE EXCLUSIVAMENTE DE UNA DIMENSIÓN O FACTOR: SU NIVEL EN LA HABILIDAD MEDIDA.
- ❖ **INDEPENDENCIA LOCAL**. INDICA QUE LOS MODELOS ASUMEN QUE LAS RESPUESTAS DE LAS PERSONAS A UN ÍTEM SON INDEPENDIENTES DE LAS RESPUESTAS A LOS OTROS ÍTEMES.
- ❖ **EXPERIENCIAS EDUCACIONALES SIMILARES** EN LOS ALUMNOS QUE RINDEN LAS PRUEBAS.
- ❖ **QUE LA PRUEBA NO HAYA SIDO APURADA**, ES DECIR, SE ASIGNE EL TIEMPO NECESARIO PARA QUE TODOS ALCANCEN A ABORDAR TODOS LOS ÍTEMES.
- ❖ **QUE NO HAYA EFECTOS DE CONTEXTO NO CONTROLADOS**.

ES DE SUMA IMPORTANCIA QUE EL MODELO SEA CAPAZ DE PREDECIR CON LA MAYOR EXACTITUD POSIBLE EL COMPORTAMIENTO DE LOS ALUMNOS FRENTE A CADA PREGUNTA.



VENTAJAS Y DESVENTAJAS DE LA TRI

- LA PRINCIPAL VENTAJA POTENCIAL SOBRE LA TC ES LA INVARIANZA DE LOS PUNTAJES DE LA PRUEBA Y DE LAS CARACTERÍSTICAS DE LAS PREGUNTAS.
- LA POSIBILIDAD DE CONSTRUIR CURVAS DE INFORMACIÓN PARA CADA ÍTEM LO QUE PERMITE LA OPTIMIZACIÓN DE LA SELECCIÓN DE LAS PREGUNTAS PARA UNA PRUEBA CON OBJETIVOS ESPECÍFICOS.
- PROPORCIONA MÉTODOS ALTERNATIVOS PARA LA DETECCIÓN DE SESGOS EN LAS PREGUNTAS. ANÁLISIS DIF.
- PROPORCIONA MÉTODOS ALTERNATIVOS PARA LA REALIZACIÓN DE PROCESOS DE EQUATING, PROCESO POR EL CUAL DOS PRUEBAS SE HACEN COMPARABLES.

SIN EMBARGO, TODAS LAS VENTAJAS ANTERIORES SE PIERDEN CUANDO NO SE CUMPLEN LOS REQUISITOS Y EN MUCHAS OCASIONES LA NATURALEZA DE LAS DISCIPLINAS MEDIDAS IMPIDE EL CUMPLIMIENTO DE LOS SUPUESTOS LO QUE HACE ACONSEJABLE NO USAR LA TRI.



ALGUNOS EJEMPLOS TÍPICOS DE NO CUMPLIMIENTO DE LOS SUPUESTOS DE LA TRI

- SE HA COMPROBADO QUE HAY DISCIPLINAS QUE SON CLARAMENTE MULTIDIMENSIONALES.
- EJEMPLOS:
 - PRUEBAS DE CIENCIAS (Hamilton et al 1997 y Nussbaum et al 1997)
 - PRUEBAS EN EL ÁREA DE LAS CIENCIAS SOCIALES
- LA MEDICIÓN DE LA COMPRESIÓN LECTORA A PARTIR DE UN TEXTO VIOLA EL SUPUESTO DE INDEPENDENCIA LOCAL. (Kolen y Brennan 1995)
- LA MEDICIÓN EN GRAN ESCALA IMPLICA QUE LOS ALUMNOS MEDIDOS HAN SIDO SOMETIDOS A DIFERENTES EXPERIENCIAS EDUCATIVAS, LO QUE TRANSGREDE EL SUPUESTO DE EXPERIENCIAS EDUCACIONALES SIMILARES, LO QUE A SU VEZ, ATENTA CONTRA EL SUPUESTO DE INVARIANZA DE LAS PREGUNTAS.





FINALMENTE, EN EL CONTEXTO DE LA SELECCIÓN , ES DECIR , EMPLEAR LA TRI CON LOS RESULTADOS DE LAS PRUEBAS DE ADMISIÓN, CONLLEVA ALGUNOS PROBLEMAS PRÁCTICOS:

- **COMPRESIÓN DE LOS PUNTAJES:** POSTULANTES CON EL MISMO NÚMERO DE RESPUESTAS CORRECTAS, ERRADAS Y OMITIDAS TENGAN DISTINTO PUNTAJE DE HABILIDAD.
- **PREPARACIÓN POR PARTE DE LOS ESTABLECIMIENTOS EDUCACIONALES DE PRUEBAS DE ENSAYO, DADO QUE NO TIENEN LA TECNOLOGÍA PARA ASIGNAR LOS PUNTAJES.**
- **DIFICULTAD POR PARTE DE QUIENES NO MANEJAN LA TRI DE VERIFICAR SI EL PROCESO HA SIDO BIEN EJECUTADO.**



ANÁLISIS DEL FUNCIONAMIENTO DIFERENCIAL DE UN ITEM DIF

CONSISTE EN ANALIZAR SI UNA PREGUNTA SE COMPORTA DE MANERA SIMILAR EN DISTINTOS GRUPOS QUE SE PUEDEN FORMAR EN LA POBLACIÓN QUE RINDE UNA PRUEBA.

LOS PARÁMETROS QUE SE ANALIZAN SON:

- DIFICULTAD**
- DISCRIMINACIÓN**

LOS GRUPOS DE ANÁLISIS GENERALMENTE RESPONDEN A CARACTERÍSTICAS QUE DE ALGÚN MODO ESTÁN PRODUCIENDO PUNTAJES DISTINTOS. POR EJEMPLO:

- GÉNERO**
- TIPO DE EDUCACIÓN**
- DEPENDENCIA DEL ESTABLECIMIENTO EDUCACIONAL**



ANÁLISIS DEL FUNCIONAMIENTO DIFERENCIAL DE UN ITEM DIF

PARA DETERMINAR SI LAS DIFERENCIAS OBSERVADAS SON PRODUCTO DE LA PREGUNTA O BIEN ES CONSECUENCIA DE QUE LOS GRUPOS ANALIZADOS SON ESTRUCTURALMENTE DISTINTOS, ES DECIR, HAN SIDO SOMETIDOS A EXPERIENCIAS EDUCATIVAS Y AMBIENTALES QUE NO SON HOMOLOGABLES, EXISTEN MÉTODOS ESTADÍSTICOS, TANTO EN LA TC COMO EN LA TRI.

EN ESTE CASO SE HABLA DE IMPACTO, NO DE SESGO DEL ITEM.

EJEMPLO: RESULTADOS EN UNA PRUEBA QUE SE APLICA A GRUPOS QUE HAN TENIDO UNA FORMACIÓN QUE SE SABE ES DIFERENTE. POR LO TANTO, NO ES LA PRUEBA LA CAUSA DE LAS DIFERENCIAS EN EL NIVEL DE LOGRO, ELLA SE ENCUENTRA EN LA FORMACIÓN DE BASE DE LOS ALUMNOS.





EN SÍNTESIS

- SE DICE QUE LOS SUPUESTOS DE LA TC SON MÁS BIEN DÉBILES , EN EL SENTIDO QUE SON GENERALES, SU FUERZA ESTÁ EN SU GENERALIDAD, SE PUEDEN APLICAR A SITUACIONES MUY VARIADAS.
- POR SU PARTE, LOS SUPUESTOS DE LA TRI SON MÁS FUERTES, MÁS RESTRICTIVOS, SE SACRIFICA GENERALIDAD PARA GANAR PRECISIÓN PREDICTIVA.
- EL PRECIO A PAGAR ES LA EXIGENCIA QUE LOS DATOS CUMPLAN SUPUESTOS MUY ESPECÍFICOS.
- EL MODELO SUBYACENTE ES REFUTABLE.
- ANTE LA ETERNA DISYUNTIVA ENTRE GENERALIDAD Y PRECISIÓN, A LA QUE TODA METODOLOGÍA CIENTÍFICA SE ENFRENTA, LA TC DA MÁS PESO A LA GENERALIDAD Y LA TRI A LA PRECISIÓN.

EN CONSECUENCIA, AMBOS ENFOQUES ESTÁN OBLIGADOS A ENTENDERSE, EN PROVECHO DE LOS USUARIOS. ESTO IMPLICA USAR LAS TEORÍAS EN FORMA COMPLEMENTARIA.



EL DEMRE HA ADOPTADO ESTA POSICIÓN DE USAR LAS DOS TEORÍAS EN FORMA COMPLEMENTARIA.

PARA ELLO SE REALIZAN LOS ANÁLISIS DE CADA PREGUNTA BAJO LA ÓPTICA DE LA TEORÍA CLÁSICA Y DE LA TEORÍA DE RESPUESTA AL ÍTEM.

LO ANTERIOR PERMITE CONTAR CON MAYOR INFORMACIÓN EN EL MOMENTO DE SELECCIONAR LAS PREGUNTAS QUE FORMARÁN PARTE DE CADA UNA DE LAS PRUEBAS, LO QUE, A SU VEZ, CONTRIBUYE A QUE LA MEDICIÓN REALIZADA SEA CADA VEZ MÁS PRECISA Y VÁLIDA.

LO ANTES EXPUESTO DA SUSTENTO AL PROPÓSITO DE LAS PRUEBAS, *“SELECCIONAR A LOS POSTULANTES QUE TIENEN LA MAYOR PROBABILIDAD DE ÉXITO EN SU PRIMER NIVEL UNIVERSITARIO”*





REFERENCIAS BIBLIOGRÁFICAS EN ESPAÑOL

Dusaillant, Francisca , *“Técnicas de Medición en Pruebas de Admisión a las Universidades”*. Centro de Estudios Públicos, Otoño 2003 N° 90

Manzi, Jorge; San Martín Ernesto, *“La Necesaria Complementariedad entre Teoría Clásica de la Medición (TCM) y Teoría de Respuesta al Item (IRT)”*. Centro de Estudios Públicos, Otoño 2003 N° 90

Muñiz, José, *“Introducción a la Teoría de Respuesta a los Items”*. Ediciones Pirámide, Madrid, 1997.

Stage, Christina, *“Teoría Clásica de Medición o Teoría de Respuesta al Item: La Experiencia Sueca”*. Centro de Estudios Públicos, Otoño 2003 N° 90.