

31961

5



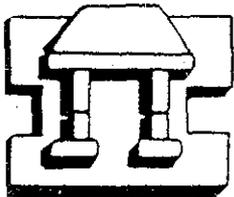
UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE ESTUDIOS SUPERIORES IZTACALA

BAREMACION DEL TEST DE INTELIGENCIA FACTOR "g" DE R B CATTELL, EN ESTUDIANTES DE NIVEL BACHILLERATO Y PRIMER AÑO DE ESTUDIOS UNIVERSITARIOS, DE ESCUELAS PUBLICAS DE LA ZONA METROPOLITANA DE LA CIUDAD DE MEXICO

T E S I S
PARA OBTENER EL GRADO DE MAESTRA EN MODIFICACION DE CONDUCTA
P R E S E N T A :
MA. DEL ROSARIO GONZALEZ VELAZQUEZ

ASESOR: DR ARTURO SILVA RODRIGUEZ



IZTACALA

LOS REYES IZTACALA, EDO DE MEXICO

2002



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

# ÍNDICE

INTRODUCCIÓN	1
CAPÍTULO 1. TEORÍAS PSICOMÉTRICAS	6
1.1 Antecedentes	6
1.2. La medida en psicología	8
1.3. Concepto de psicometría	12
1.4. teorías psicométrica	14
1.4.1 Teoría Clásica de los test (TCT)	16
1.4.1.1 Supuestos básicos de la TCT	20
1.4.1.2 Confiabilidad	23
1.4.1.3 Métodos de la TCT para estimar la Confiabilidad	24
1.4.1.4 Ventajas y desventajas de los métodos	25
1.4.1.5 Limitaciones de la TCT	27
1.4.2 Teoría de la Generabilizabilidad (TG)	28
1.4.2.1 Marco conceptual	30
1.4.2.2 Estudios de generabilizabilidad	32
1.4.3 Teoría de la Respuesta al ítem	35
1.4.3.1 Antecedentes históricos	37
1.4.3.2 Definición	40
1.4.3.3 Curva característica de los ítems	43
1.4.3.4 Modelos paramétricos	44
CAPÍTULO 2. LA INTELIGENCIA Y SU MEDICIÓN	50
2.1 Origen de la inteligencia	51-
2.2 Cómo se define la inteligencia	54
2.2.1 Mitos y verdades acerca de la inteligencia	60
2.3 Teorías que explican la inteligencia	62
2.3.1 Teorías psicométricas o diferenciales de la inteligencia	66

2.3.1.1	Modelos monolíticos o unifactoriales de la inteligencia	70
2.3.1.2	Modelos factoriales de la inteligencia	71
2.3.2	Teorías cognitivas o del procesamiento de la información sobre la inteligencia humana	74
2.3.2.1	Teorías múltiple de la inteligencia	76
2.3.3	Teorías sobre inteligencia emocional	80
2.3.4	Fortalezas y debilidades de los diversos modelos explicativos de la inteligencia	81
2.3.5	Los test de inteligencia: origen, tipos y características	85
2.3.5.1	Definiciones	85
2.3.5.2	Origen de los test de inteligencia	86
2.3.5.3	El factor "g"	96
2.3.5.4	Construcción del factor "g"	101
2.3.5.5	Características de los test de inteligencia	102
2.3.5.6	Clasificación de los test	112

### CAPITULO 3. BAREMACIÓN DEL TEST DE FATOR "g" EN SU ESCALA 3 DE R.B. CATTELL Y A.S.K. CATTELL EN LA CIUDAD DE MÉXICO 118

3.1	Planteamiento del problema	118
3.2	Objetivo general	122
3.3	Hipótesis de trabajo	123
3.4	Método	123

3.4.1	Sujetos	123
3.4.2	Instrumentos de medida	127

3.5	Análisis Estadístico	129
-----	----------------------	-----

3.5.1	Procedimiento seguido para la baremación	132
-------	--	-----

3.6	Resultados	134
-----	------------	-----

### CONCLUSIONES 139

### REFERENCIAS 144

## INTRODUCCIÓN

La evaluación psicológica se considera como una actividad sistemática que se encuentra integrada a tres grandes áreas de actividad: la clínica la educativa y la industrial; y tiene por objetivo proporcionar tanto a los psicólogos como a los profesionales de la educación la máxima información posible sobre él o los sujetos con quienes trabaja para así poder mejorar tanto en lo individual como en lo colectivo los procesos de integración y desarrollo de los individuos en dichos ámbitos.

Una gran parte del trabajo de la evaluación psicológica, se fundamenta en la utilización de pruebas psicométricas, que como instrumentos técnicos está sujeto a muchas expectativas en cuanto a sus alcances y limitaciones, así como a problemas en su aplicación a la investigación y al diagnóstico. Una de las disciplinas que se ha abocado a la resolución de estos problemas es la psicometría, que por mucho tiempo se ha identificado con la construcción de instrumentos (tests) que sirvan para asignar precisamente números a atributos o conductas de las personas. En los últimos años ha adquirido una gran importancia ya que remite al establecimiento de un pensamiento objetivo y a la apreciación de los fenómenos psíquicos por medio de su cuantificación, valorando su adecuado uso dentro de la psicología aplicada (Serrano. 1984).

Históricamente los test se construyeron como un método de investigación en psicología diferencial, vislumbrándose desde sus primeras instrumentaciones las aplicaciones prácticas que podían tener sin perder con ello, las aportaciones

que por sí mismos dan al desarrollo de la psicología teórica (Pichot, 1979). El análisis factorial facilitó en gran medida el desarrollo de los test, permitiendo que dicho método no se limitara a una sola escuela en particular, sino que se ampliara a otras corrientes de pensamiento como el asociacionismo, el conductismo, la psicología gestáltica, el psicoanálisis entre otras. quienes a su vez dieron origen a otros test extrayendo de su aplicación la confirmación o invalidación de sus hipótesis base.

Actualmente existe una gran variedad de pruebas que intentan medir diversos tipos de comportamiento, entre ellas se encuentran los llamados test de inteligencia cuyo uso se ha extendido ampliamente en diversas instituciones, sobre todo en las educativas, en donde son utilizada principalmente para obtener diagnósticos sobre el posible rendimiento académico de los alumnos. La inteligencia se considera como uno de los factores esenciales que determinan en gran medida el éxito o el fracaso escolar, de hecho se puede decir que el origen de estas pruebas tiene que ver con la valoración de dicha facultad y también con la existencia de una gran número de teorías que intentan explicarla.

La inteligencia general figura entre los conceptos más elusivos que ha sido definido tanto por expertos como por los no expertos. De aquí han derivado teorías explícitas e implícitas de la inteligencia, considerando a las primeras como aquellas descripciones formales de la inteligencia formuladas por psicólogos; mientras que las teorías implícitas son las nociones informales que tenemos o llevamos con nosotros sobre la inteligencia. Las teorías explícitas conforman la base de la mayor parte de las investigaciones empíricas, de aquí su importancia ya

que son éstas las que dan fundamento a la elaboración de instrumentos como el que hoy se pretende analizar

Las teorías explícitas de la inteligencia están basadas y/o comprobadas sobre datos recogidos de personas realizando tareas o ejercicios que suponen miden el funcionamiento inteligente. De estas investigaciones Sternberg (1990) identifica dos vertientes teóricas de mayor influencia; las teorías diferenciales y las teorías cognoscitivas, (ambas vertientes serán desarrolladas con puntualidad en el capítulo dos de este trabajo) Dentro de las primeras se encuentran las teorías del factor general o "g" como frecuentemente se le denomina, y sugiere que la naturaleza de las diferencias individuales puede ser entendida en término de los niveles de energía mental que el individuo utiliza en la realización de una tarea intelectual, y/o en términos de las aptitudes que el sujeto utiliza para realizar la tarea a partir de tres principios cualitativos: percepción de la experiencia, educación de las relaciones y educación de los elementos correlacionados.

Ubicado dentro de este contexto encontramos los trabajos realizados por R.B. Cattell (1994), quien analiza el problema de la medición de la inteligencia desde lo que él denomina *inteligencia fluida* e *inteligencia cristalizada*; la "primera representa la capacidad básica y la segunda representa las capacidades adquiridas a través del aprendizaje, la práctica y la exposición a la educación" (Carroll, 1982, p.118). La construcción del factor "g" lleva al autor y a otros investigadores a la búsqueda de una medida más pura de la inteligencia, centrando su atención en la creación de instrumentos que le permitieran eliminar los efectos culturales o aquellos ligados a la información, manteniendo los más puramente procesuales, aportando como resultado que las aptitudes vinculadas al

procesamiento lógico, a la inducción y a la generalización constituyen el núcleo de los instrumentos de factor "g", depurados de elementos verbales (muy ligados al origen sociocultural) y minimizando el acceso a la memoria con el ánimo de convertirlos en medidas libres de cultura (Genovord, Rosello, Beltan, et al 1995). Esta condición posiblemente es la que ha permitido que este tipo de pruebas tengan aún vigencia y sigan siendo ampliamente utilizadas, ya que constituyen un instrumento de gran utilidad para evaluar la inteligencia.

El test de factor "g" de Cattell, constituye un instrumento de gran utilidad para evaluar la inteligencia, ya que puede ser aplicado a una serie de situaciones en las que resulta muy recomendable su utilización. Originalmente existen tres escalas (1, 2, y 3) adecuadas para varios niveles de edad; las Escala 2 y 3, destinadas a poblaciones de adolescentes y adultos, y la Escala 1, aplicable a niños con edades comprendidas entre cuatro y ocho años o a adultos con deficiencia mental, a diferencia de las otras dos escalas, ésta "no se puede considerar completamente libre de influencias culturales, ya que resulta complejo elaborar un número suficiente de elementos con material de tipo perceptivo, capaz de atraer y mantener la atención de los niños de poca edad y cumplir los demás requisitos que exige el ámbito de aplicación" (Cattell, 1989 p.4) Por su parte las Escalas 2 y 3 presentan cuatro subtes los cuales presentan elementos en forma gráfica y prácticamente sin ningún contenido cultural. La Escala 2 está diseñada para ser aplicada a escolares de edades comprendidas entre 8 y 14 años y adultos de escasa formación académica; mientras la Escala 3 se dirige a estudiantes en últimos cursos de enseñanza media y universitarios, y en general adultos con un nivel cultural relativamente elevado

Debido a que esta prueba forma parte del acervo de instrumentos que son utilizados en la Unidad de Evaluación psicológica (UEPI) de la Facultad de Estudios Superiores Iztacala (FES-I), los baremos que en ella se incluyen corresponden a grupos específicos del país de procedencia. En el caso de este test su adaptación se realizó en España, país que cuenta con una fuerte producción de materiales psicométricos que son exportados a diversos países entre ellos México, por lo tanto los baremos incluidos en éstos son utilizados para realizar las interpretaciones correspondientes, lo que tiende a sesgar la información obtenida sobre el rendimiento de nuestros grupos específicos de población, de aquí que se haya decidido baremar la prueba de Cattell con la finalidad de continuar con su utilización de forma más confiable. Como ya se mencionó, es una de las más usadas en la UEPI para obtener el CI., de nuestros estudiantes y usuarios.

La presentación de este trabajo la iniciaremos con el desarrollo de las teorías psicométricas que dan sustento a la construcción de instrumentos que pretenden medir las capacidades humanas. La intención del primer capítulo es mostrar cómo se aplican estos principios, a partir del actual conocimiento científico en el cual se fundamentan los test psicológicos, para poner posteriormente el acento en los test de inteligencia, (capítulo dos), los cuales se encuentran apuntalados por una extensa gama de planteamientos teórico-metodológicos y técnicos; para finalmente concretarnos a la baremación del Test de factor "g" en su Escala 3 de R.B. Cattell y A.K.S. Cattell, en estudiantes de Educación Media Superior y Superior de escuelas públicas de la zona metropolitana de la ciudad de México.

# CAPÍTULO 1

## TEORÍAS PSICOMÉTRICAS

La búsqueda de explicaciones científicas de aquellos fenómenos que rodean al ser humano es una tarea a la cual se han abocado un sinnúmero de científicos e investigadores, cuyo objetivo primordial es producir nuevos conocimientos, conocimientos que puedan ser difundidos a través de informes científicos, monografías, revistas, libros, etcétera, y que a su vez éstos puedan ser aprovechados por aquellos usuarios que se dedican o centran su atención en estudios especiales. Tal es el caso de los conocimientos generados por la Psicología y la Pedagogía. Algunos de los productos creados por estas disciplinas, son objetos físicos concretos como los conocidos test mentales, los cuales son “un producto concreto de la subdisciplina y tecnología conocida como psicometría, que es análogo a los productos que incorporan conocimientos de las ciencias físicas y biológicas” (Carroll, 1987).

¿Pero de dónde proviene la psicometría?

### 1.1 Antecedentes

Un primer antecedente lo encontramos en los estudios realizados por el eminente astrónomo alemán Federico Guillermo Bassel (n Miden, 1784, m. Königsberd, 1886), quien llegó al conocimiento de la astronomía a través del estudio de la navegación. Con la construcción de un sextante, realizó sus primeras observaciones; siendo uno de esos pocos hombres que se fijan en aquellas cosas

que la mayoría de las personas pasan desapercibidas, descubrió en una historia del observatorio astronómico de Greenoich un incidente registrado allí; un joven auxiliar había sido despedido por anotar el paso aparente de estrellas a través de una línea en el campo de visión del telescopio, con un segundo de desfase respecto al tiempo anotado por su maestro, esto llevó al investigador a preguntarse a qué se debía esa diferencia que le había costado el puesto al joven, por lo que empezó a poner a prueba a sus colegas astrónomos y descubrió que había una variación considerable entre los individuos en la rapidez con que reaccionaban ante un estímulo visual. A esta variación la bautizó con el nombre de ecuación personal. La importancia de esto consistió en demostrar que los procesos mentales, uno de los procesos más difíciles de abordar por la ciencia, podrían ser cuantificados y sometidos a medida objetiva. De esta forma daba sus primeros pasos a la psicometría (Sánchez, 2001).

La necesidad de obtener cuantitativamente los atributos o características de un segmento de la conducta, en este caso de los procesos de pensamiento, ya habían sido estudiados desde la antigüedad por pensadores como Aristóteles y Platón, pero no es sino hasta la primera mitad del siglo XX que los estudios formales con bases científicas para la comprensión y predicción del comportamiento humano, adquieren su verdadera significancia. Sin embargo, no pueden dejarse de lado los primeros estudios que se llevaron a cabo alrededor de 1830 en el campo de la "psicología y las matemáticas destinados a lograr la medición mental, distinguiéndose tres corrientes de influencia: la *francesa* con Esquirol, Pineal, Itard, Seguin, Charcot y Ribot; la *alemana* con Weber, Fechner y Wunt, y la que implantó Laplace en el campo de las matemáticas, quien junto con Gauss, demostró la aplicabilidad y utilidad de la curva normal como instrumento

matemático aplicable a la medición de las características del ser Humano"  
(Morales, 2000 p. 13)

Por tanto los antecedentes de la psicometría los encontramos en el afán de las ciencias empíricas de predecir, describir y explicar los fenómenos de su área específica. Se busca recoger datos, compararlos para posteriormente fundamentar teorías<sup>1</sup>. Si partimos del hecho de que la ciencia psicológica se define a través de teorías psicológicas, -las cuales se conceptualizan como un sistema de proposiciones referidos a un conjunto de constructos, que sirven para describir, explicar y predecir algunos aspectos limitados de su dominio, (Martínez, 1995)- éstas buscarán su fundamento para el tema que nos ocupa; la **medida**, componente central, basado en instrumentos matemáticos aplicables a la ciencia, por lo que la psicología como una ciencia empírica, está en busca de dicho fundamento el cual se traduce en la psicometría.

Antes de pasar a una definición más precisa sobre la psicometría, es necesario detenernos un momento en la explicación de lo que se entiende por medida.

## 1.2 La medida en psicología

En un sentido más estricto, se entendería que el concepto de medición, implica "la presencia de algún tipo de escala dividida en unidades iguales, y en la cual existe

---

<sup>1</sup> El término teoría es utilizado en la literatura bajo dos acepciones principales; en sentido amplio se trata de una ciencia del conocimiento general, a diferencia de la actividad práctica del hombre. La teoría aparece también sobre la base de la práctica, pero no simplemente generaliza la experiencia práctica, sino que va más allá. Descubre nuevos nexos y relaciones del objeto y ayuda con ello mismo a la práctica a asimilarlo con éxito. En sentido estricto es el conocimiento que tiene una forma rigurosamente determinada (Blauberg, 1986)

una calificación de cero que corresponda a nada de lo que se está midiendo" (Adkins, 1996, p 24). Así sucede por ejemplo con el metro en donde un centímetro es igual para cualquier parte de la escala, y donde cero centímetros tiene el significado racional de correspondencia con la ausencia total de longitud. por tanto podemos confiar en la validez de dicho instrumento. ya que las unidades que lo conforman son iguales y el origen o punto cero de la escala tenga alguna *significación racional, que corresponda a la que le damos a la nada en la realidad física*, sin embargo, estos mismos principios no pueden ser aplicados a los instrumentos de medición psicológica o educativa, ya que una calificación "cero" raramente corresponde a la ausencia absoluta de conocimientos o habilidad que se desee medir, así, una de las funciones de "toda teoría consiste en emplear constructos hipotéticos, y las desviaciones de medida de esos valores hipotéticos constituyen la explicación de los eventos paradójicos". (Morales, 2000, p.10)

La medición en psicología como en cualquier otra ciencia se inicia con la identificación de elementos del mundo real. Con los elementos o construcciones de un sistema lógico-abstracto al cual se le llama modelo. Una vez definido el elemento del mundo real que se pretenda conocer, se puede proceder a la medición, para lo cual deben seguirse los siguientes pasos: "1) identificar objeto o sujeto a medir; 2) identificar la propiedad o conducta que se quiere medir en el objeto o sujeto y el instrumento; 3) aplicar una regla de asignación numérica por la que se asigna un número a la propiedad de la unidad experimental (objeto o sujeto) que se está midiendo" (Martínez, 1995, p 23)

Se puede decir entonces que las designaciones hechas al objeto o sujeto en cuestión, constituyen una medida de la que se deriva el concepto de medición,

que de acuerdo con Nunnally (1970) consiste en un conjunto de normas para asignar números a los objetos de modo tal que estos números representan cantidades de atributos. Otras definiciones las encontramos con "Lord y Novick (1968) quienes la definen como un procedimiento para la asignación de número (puntuaciones, medidas) a propiedades específicas de unidades experimentales, de modo que las caractericen y preserven las relaciones específicas en el dominio comportamental (p.17) Para Stevens (1951, p.22) la medida es la asignación de número a objetos o sucesos según ciertas reglas" (en Martínez, 1995, p. 23-24).

Cualquiera que sea la definición que se plantee, lo cierto es que la validez o utilidad de una medida va a depender siempre del carácter de los datos empíricos que se tengan. Se requiere que la medida concuerde con un conjunto de axiomas de medición; un modelo.

Los científicos establecen medidas, reglas para la cuantificación de atributos de objetos reales; como ya se mencionó toman en préstamo sistemas matemáticos para analizar las relaciones internas de los datos que han obtenido de una medición y para relacionar diferentes medidas entre si, lo fundamental es la elección del método de medición, y es que la justificación de su uso exige que en el mundo real se encuentre una serie fértil de relaciones lícitas entre variables, (Nunnally, 1991) y como en todo estudio científico, en el campo de la psicología debe haber diversas formas de medición; lo que deriva al asignar números a los atributos de los objetos a lo que Morales (2000) denomina: niveles de medición, o dicho con más propiedad "escalas de medición, que son elementos que se emplean frecuentemente de la siguiente manera: a) escalas nominales (o de categorización); b) escalar ordinales; c) escalas intervalales y d) escalas de razón"

(p 13). De estas cuatro, sólo se pueden utilizar las tres primeras ya que la última por partir de cero absoluto, la mayoría de las veces no es posible emplearla en la medición psicológica, debido a que en psicología siempre se inicia de "algo", y ese algo es lo que lleva a la búsqueda de explicaciones de los fenómenos empíricos de su área de interés, con la finalidad de establecer principios generales para explicar y predecir dichos fenómenos; por lo tanto, como ciencia debe recoger y comparar los datos para establecer correlaciones, ecuaciones, etc., que le permitan fundamentar sus teorías. En donde la medida se presenta como un componente fundamental de ese proceso y es lo que permite que los instrumentos de las matemáticas puedan aplicarse a la ciencia (Torgerson en Martínez, 1995). No es extraño por tanto que una materia importante de esta ciencia sea la psicometría, la cual se define a continuación:

### 1.3 Concepto de psicometría

La psicometría se fundamenta en las llamadas ciencias empíricas que tienen como meta describir, predecir y explicar los fenómenos empíricos, y en donde la medida es el componente fundamental en dicho proceso, por ello la psicometría se define como aquella rama de la matemática que tiene por objeto la medición de la actividad mental, la elaboración de diferentes instrumentos de medida (test, cuestionarios, escalas, etc.) y el posterior tratamiento matemático de los datos obtenidos. Su origen se encuentra en los estudios realizados en el campo de las diferencias sensoriales por investigadores como Wunt, Weber, Fechner, Galton, Pearson y Spearman, principalmente (Diccionario de Ciencias de la Educación, 1985).

Para Martínez (1995, p. 21), "La psicometría aglutina todo el conjunto de modelos formales que posibilitan la medición de variables psicológicas, centrándose en las condiciones que permiten llevar a cabo todo el proceso de medición en psicología y en establecer las bases para que estos procesos se realicen en forma adecuada (...) Su objetivo es desarrollar modelos de índole cuantitativa para la transformación de los hechos en datos y proporcionar métodos idóneos para la aplicación de esos modelos, con objeto de asignar valores numéricos generalmente a los sujetos, sobre la base de sus respuestas y/o a los estímulos presentes en la situación".

Yela (1968) sostiene que la Psicometría se ocupa de todas las medidas en el campo psicológico, habiéndose desarrollado a través de dos ramas un tanto independientes: la de los métodos psicofísicos y la del método de tests

Para Nunnally (1973) es la metodología que se refiere al desarrollo y empleo de las técnicas de medida en todos los aspectos de la Psicología.

Rivas (1979) llega a la conclusión de que existe "un consenso general para asignar a la Psicometría la tarea de medir comportamientos observables y de naturaleza psicológica".

Para Maciá (1982) la Psicometría es una disciplina encuadrada dentro del área de la Ciencia Psicológica, con una doble función: teórica, en cuanto tiene como misión el estudio de la medida y sus posibilidades en el campo psicológico, y práctica, en cuanto aplica esos conocimientos al estudio de aspectos psicológicos concretos, tanto en lo que se refiere a como obtener las medidas (métodos) como a con qué obtenerlas (instrumentos)

Muñiz (1992) define en términos generales la Psicometría como el conjunto de métodos, técnicas y teorías implicadas en la medición de variables psicológicas, siendo su característica específica el énfasis y especialización en aquellas propiedades métricas exigibles a las mediciones psicológicas independientemente del campo sustantivo de aplicación y de los instrumentos utilizados

Barbero (1994) considera que la Psicometría es una disciplina metodológica, dentro del área de la Psicología, cuya tarea fundamental es la medición o cuantificación de las manifestaciones de la conducta con todas las implicaciones que ello conlleva, tanto teóricas (posibilidades y criterios) como prácticas (cómo y con qué se mide)

De acuerdo con estas definiciones, la Psicometría deberá ocuparse no sólo de las *justificación y la legitimación de la medición psicológica*, en sus diversos campos, dándole una fundamentación teórica cuyos principios constituyen la *Teoría de la Medición*, o psicométricas sino de las implicaciones prácticas y aplicadas que dicha medición conlleva, desarrollando los métodos necesarios que nos indiquen, en cada caso concreto, cómo se debe llevar a cabo la cuantificación, construyendo los *instrumentos* necesarios para poder efectuarlos (Barbero, 1994, p. 4). Es claro que esos instrumentos suelen tener una influencia significativa en las personas tanto en sus aspectos personales como profesionales, por lo que se considera fundamental que el usuario de los mismos, conozca los procedimientos en los que se basan sus puntuaciones, la información que proporcionan y las limitaciones que entrañan. Si la finalidad de la psicometría es aportar soluciones al problema de la medida en cualquier proceso de investigación psicológica, debe incluir aspectos tanto teóricos como de carácter práctico.

En su perspectiva teórica: incluye las teorías que tratan de las medidas en psicología, encargándose de describirlas, categorizarlas, evaluar su utilidad y precisión, así como la búsqueda de nuevos métodos, teorías y modelos matemáticos que permitan mejores instrumentos de medida. Veamos en que consisten.

#### **1.4 Teorías psicométricas**

La teoría psicométrica recibió su primer impulso de los estudios que sobre la psicofísica realizó Fechner, quien no se conformó con generalizar la psicofísica más allá de los simples estudios de juicio humano, sino que creía firmemente que en psicología es posible y factible la medición exacta. Sus estudios fueron un aliciente para la búsqueda de métodos precisos en la psicología general.

Posterior a los estudios de Fechner, le siguieron en orden de importancia los de Galton y Pearson, quienes desarrollaron técnicas estadísticas para analizar los datos obtenidos con los métodos de medición e idearon procedimientos para analizar las correlaciones entre los diferentes métodos. Siguiendo la misma línea Spearman fue otro de los investigadores que realizaron grandes contribuciones a la teoría psicométrica, sobre todo en lo que se refiere a la construcción de los test, con el desarrollo de la teoría de la confiabilidad y muchos aspectos del análisis de correlación y el análisis factorial (Nunnally, 1970)

El desarrollo posterior de la teoría psicométrica se dio en los Estados Unidos, presentándose en la actualidad como una teoría tan polifacética y compleja que se ha constituido en un subcampo especial de la psicología. Campo

que engloba aquellas teorías que "se centran principalmente en la cuestión de qué constelaciones estructurales de los procesos, contenidos y representaciones dan lugar a los modelos estables de diferencias individuales ( ..) por lo que el enfoque es útil para describir cómo se agrupan los procesos, contenidos y representaciones para formar fuentes relativamente estables de diferencias individuales" (Sternberg, 1986, p.26-27).

De acuerdo con Carroll (1987) para que una teoría explique lo que los test mentales miden y en qué medida lo miden bien, tiene que exponer de modo satisfactorio al menos los siguientes puntos:

- 1) *Hasta qué punto las tareas individuales o en grupo, pueden ofrecer u ofrecen medidas fiables de los comportamientos característicos de los individuos que ejecutan dichas tareas.* En proporción en que las medidas derivadas de las tareas sean realmente fiables, podrá considerarse que los individuos pueden ser dispuestos u ordenados en una escala o continuo que refleje los diferentes grados de competencia respecto a la clase o clases de la tarea representada por aquellas cuya ejecución ha sido observada
- 2) *Hasta qué punto la ejecución con éxito de cada tarea o grupo de tareas refleja cierto tipo particular de competencia o potencial cognitivo.* La teoría deber ofrecer una clasificación satisfactoria de los diversos tipos de competencia que pueden ser observadas, y que le corresponda una escala o continuo que contenga los diferentes grados de competencia.
- 3) *La fuente o causa de la diferencias individuales observadas respecto a cualquier escala de competencia cognitiva identificada como tal.* (p. 104)

Debido a que la teoría psicométrica ha sido muy matematizada, para especializarse en ella se requiere de un alto grado de conocimientos matemáticos, sin embargo ello no significa que sus principios más importantes puedan ser

comprendidos por quien apenas se inicia en el estudio de este campo, por ello en este capítulo se tratará de exponer los rasgos más generales de las teorías que por el modelo estadístico en que se inscriben, se clasifican en; Teoría Clásica de los test, Teoría de la generalizabilidad y Teoría de la Respuesta al Ítem. Poniendo el énfasis en la primera, ya que es de ella de donde se obtiene el fundamento teórico-metodológico del test de factor "g" de Cattell, objeto de estudio de esta investigación.

#### **1.4.1 Teoría Clásica de los Test (TCT).**

Como de ha visto hasta el momento, los primeros fundamentos de la teoría de los test, se dieron a la par con el primer movimiento de los test mentales. Durante este periodo la teoría de los test se ocupaba principalmente de la confiabilidad de las mediciones "que ya habían sido cuantificadas (por ejemplo, las puntuaciones de 'número correcto' en los test mentales y los 'coeficientes de inteligencia' ); la preocupación de cómo tuvo lugar esta cuantificación, tan sólo se expresaría con el desarrollo de diversas técnicas de 'análisis de ítems' , destinados a asegurar que todos los ítems o tareas que intervienen en la obtención de una medición dada tendrían en promedio a coincidir entre sí, estableciendo una eficaz contribución a la confiabilidad de la medición, la confiabilidad fue destinada como la medida en que las evaluaciones observadas reflejan un resultado subyacente 'verdadero' por oposición al error. Este enfoque es conocido como "teoría clásica de los test" (Novick, 1966; Stanley, 1971, en Carroll, 1987, p. 105)

Se puede considerar que la TCT, es un modelo relativamente sencillo que fundamentalmente describe como los errores de medida pueden influir en las puntuaciones observadas de un individuo en un test.

Los primeros modelos formales para la estimación de los errores inherentes a la medición psicológica fueron propuestos a principios del siglo XX por Spearman (1904) su propuesta conocida actualmente como modelo lineal clásico, "postula que la puntuación empírica de un sujeto ( $x$ ), estaría compuesta por dos componentes aditivos, la verdadera puntuación de la persona en la prueba ( $v$ ), y un error aleatorio ( $e$ ):  $x = v + e$ " (Muñiz, 2001, p.2), y en donde el problema consiste precisamente en cómo estimar el verdadero valor del error ( $e$ ) Por ello Spearman añadió a su modelo tres supuestos y una definición a partir de la cual se derivan la mayoría de las formulas y conceptos de la TCT. El modelo representado matemáticamente es el siguiente:

<i>Modelo Lineal Clásico</i>
<b>Modelo <math>X = V + e</math></b>
<b>Supuestos: <math>V = E(X)</math></b>
$\Delta_{ve} = 0$
$\Delta_{(e_j, e_k)} = 0$
<b>Definición: Dos test <math>j, k</math> se consideran paralelos si: <math>V_j = V_k</math> y <math>\Phi_{e_j} = \Phi_{e_k}</math></b>

Tomado de Muñiz (2001 p. 2)

Antes de pasar a la explicación del modelo lineal clásico se considera conveniente definir algunos conceptos importantes:

*Puntuación Verdadera*; se define como la puntuación que realmente le corresponde a un sujeto en una variable determinada. Esta puntuación es la que se intenta obtener con cualquier test o instrumento de medida. Sin embargo, puesto que los instrumentos de los que se dispone para medir no son instrumentos perfectos, la *puntuación empírica u observada* a través de estos instrumentos se supone que está afectada por un error. Es decir, cuando se mide algo, se comete un error puesto que no existen instrumentos de medida que sean perfectos. Este error que se comete al medir es lo que se denomina *error de medida*. La medida empírica u observada es la que se obtiene de una observación directa e inmediata del instrumento (García, 1993, p 20).

Es importante aclarar, que en psicología se presenta un problema en cuanto a la variable que se pretende medir y es que; no son directamente observables como es el caso de los fenómenos físicos, sino que son variables latentes (es decir, que sólo se mide lo que se ve a través de la manifestación de la conducta) y por lo tanto la puntuación que se obtenga de un sujeto en esa variable tampoco será verdadera, no obstante teóricamente se puede utilizar el concepto de puntuación verdadera, además de que sus implicaciones pueden ser verificables en la práctica diaria.

Por consiguiente, "la TCT esta basada en una definición de puntuaciones verdaderas, matemáticamente aceptable y conceptualmente utilizable y en ciertos

supuestos básicos que relacionan la puntuación verdadera y el error de medida" (Lord y Novock, en García. 1993 p.21).

Considerando los elementos expuestos hasta el momento, se podría denominar la TCT; "al conjunto de principios teóricos y métodos cuantitativos derivados de ellos, que fundamentan la construcción, aplicación, validación e interpretación de los distintos tipos de test, y que permiten derivar escalas estandarizadas aplicables a una población" (Díaz, 1996, p.54)

En el desarrollo de la TCT se pueden identificar tres grandes etapas desarrolladas por Cattell (en Díaz, 1996), que se identifican por su objetivo, así como por los métodos cuantitativos y tipo de análisis teórico que utilizan:

- I. **Etap. Intemetría**, se caracterizaba principalmente por la construcción de pruebas conformadas por reactivos cuyas propiedades estadísticas eran el centro de atención principal. Los test se consideraban como el producto de la integración de un conjunto de reactivos, cuyas propiedades estadísticas tenían que determinarse antes de que se les incluyera en esa prueba particular, lo que propició que el concepto de confiabilidad destacara como principal virtud de la escala y se media a partir de la correlación entre los reactivos individuales y el instrumento en su conjunto. (...) La intemetría hizo contribuciones valiosas a la psicología debido al énfasis que puso en el análisis del error, entre sus contribuciones más importante se encuentra el concepto de error estándar de medición, y a varios principios básicos de escalamiento; así como al uso generalizado de la curva normal, las pruebas de significación, el uso de correlaciones múltiples y la fórmula de atenuación, entre otras.
- II. **Etap. Psicometría estructural**. Se caracteriza por el uso de las nuevas herramientas estadísticas tales como el análisis factorial con sus variantes técnicas R-, dR-, P-, como un medio para encontrar la "estructura natural" de las habilidades en el entorno de los factores culturales, la dotación genética, la personalidad, los rasgos, los motivos dinámicos y las dimensiones que dan lugar a la acción y al comportamiento (...). La construcción de los test verdaderamente estructurados se produjo como resultado de esta etapa, por lo menos en lo que respecta a las fases iniciales del desarrollo de los mismos.

- III. Etapa. **Funcional.** Es aquella que trasciende a las aplicaciones inmediatas y simplistas que identificaban a las estadísticas con factores conductuales y profundizan en las leyes y formulaciones conceptuales del comportamiento ( ) Esas leyes según Cattell se refiere a las relaciones sistemáticas y consistentes obtenidas de los estudios empíricos sobre el desarrollo, en el conocimiento acerca de los rasgos determinados de manera hereditaria, de los rasgos modificables por las experiencias y el aprendizaje y de la modulación de los estados psicológicos producidos por las relaciones psicológicas. De esta manera, el uso funcional de los test es sólo posible en el psicólogo formado con base en las modernas teorías de la personalidad, del aprendizaje y sustentada en las técnicas de medición y experimentación. (p. 54-55)

En resumen el desarrollo de las TCT, para este autor ha precedido de etapas orientadas básicamente de manera pragmática hacia el desarrollo de los test y la validación de reactivos, hacia etapas más elaboradas en lo conceptual, en que los tests se derivan de teorías del comportamiento más articuladas y donde cada reactivo tiene un significado conceptual definido en el entorno teórico particular (Díaz, 1996).

#### 1.4.1.1 Supuestos básicos de la teoría clásica de los test.

A continuación se exponen los aspectos fundamentales del modelo lineal clásico, su formulación, asunciones y deducciones más significativas.

Formula del Modelo

$$\text{Modelo } X = V + e$$

**Supuestos:**

1.  $V = E(X)$
2.  $p(v, e) = 0$
3.  $p(e_j, e_k) = 0$

**Definiciones:** Dos test,  $j$  y  $k$ , se denominan paralelos si la varianza de los errores es la misma en ambos [ $\sigma(e_j) = \sigma(e_k)$ ] y también lo son las puntuaciones verdaderas de los sujetos ( $V_j = V_k$ ).

Tomada de Muñiz (1992)

El modelo establece que la puntuación empírica ( $X$ ) que obtiene un sujeto en un test es igual a la suma de dos componentes: la puntuación verdadera ( $V$ ) del sujeto en ese test y el error de medida ( $e$ ) cometido en la medición. Es razonable pensar que la puntuación empírica que obtiene un sujeto cuando se le aplica un test en un momento dado no coincida exactamente con su verdadera puntuación en ese test, pues en ese momento puntual el sujeto está afectado por múltiples factores de difícil control que inciden en su conducta, por lo que las puntuaciones obtenidas pueden ser más bajas de lo que realmente le correspondería, si son favorables las circunstancias, se verá beneficiado con puntuaciones superiores. Por lo tanto cuando a un sujeto se le aplica un test, no hay manera de saber si la puntuación empírica obtenida es verdadera, es decir, habrá que estimarla basándonos en los supuestos de un modelo, veamos sus fundamentos:

### **Supuesto 1**

La puntuación verdadera ( $V$ ) es la esperanza matemática de la empírica:  $V = E(V)$ , donde  $X$  es la variable aleatoria "puntuación empírica del sujeto". Este primer supuesto constituye en realidad una definición de la puntuación verdadera. Si aplicamos un test  $N$  veces a un mismo sujeto, la puntuación verdadera obtenida por éste, será la media aritmética de las puntuaciones empíricas obtenidas en las "infinitas aplicaciones". La puntuación verdadera es por tanto un concepto matemático. Es importante aclarar que la puntuación empírica de un test es una muestra empírica de conducta que si reúne ciertos requisitos de medida, y bajo ciertos supuestos, permite hacer inferencias probabilísticas fundadas, por lo que

las puntuaciones verdaderas no deben ser consideradas como algo mágico y estático, propiedad de los sujetos y que determina su conducta

### **Supuesto 2**

Se asume que no existe correlación entre las puntuaciones verdaderas de los sujetos en un test y sus respectivos errores de medida:  $\rho(v, e) = 0$ . No hay razón para pensar que el tamaño de los errores vaya sistemáticamente asociado al tamaño de las puntuaciones verdaderas.

### **Supuesto 3**

Los errores de medida de los sujetos en un test no correlacionan con sus errores de medida en otro test distinto:  $\rho(e_j, e_k) = 0$ . Si se aplican correctamente los test los errores serán aleatorios en cada ocasión, no existiendo razón a priori para que covaríen sistemáticamente unos con otros.

Ninguno de los supuestos de este modelo, son comprobables empíricamente de un modo directo tal y como han sido expresados, por lo tanto, habrá que hacer deducciones que si se puedan contrastar y confirmen y falseen el modelo (Muñiz, 1992, p. 27-28)

A continuación centraremos nuestra atención en la confiabilidad por la importancia que tiene en cuanto a la determinación del grado de error en los instrumentos de medida, fundamento central las teorías psicométricas. Y en especial de la TCT.

### **1.4.1.2 Confiabilidad**

Se considera que las mediciones psicológicas, al igual que en las de cualquier otra ciencia deben ser fiables, es decir, estar libres de errores de medida. En el caso de los test, se considera que son fiables cuando las medidas que se hacen con él carecen de errores de medida, que son consistentes. Dicha consistencia se obtiene si cada vez que se aplica un test a los mismos sujetos da el mismo resultado. Aunque dicho razonamiento pueda resultar simple y lógico, no lo es tanto ya que a diferencia de la medición de los fenómenos físicos, los seres humanos pueden presentar cambios en el comportamiento a la hora de aplicar el test, por lo que resulta problemático saber con seguridad si la "inestabilidad observada en las mediciones de debe a la imprecisión del instrumento o a los cambios legítimos operados por los sujetos" (Muñiz, 1992, p. 31)

Por tanto los errores de medida de los que se ocupa la confiabilidad son aquellos no sometidos a control e inevitables en todo proceso de medir, sea físico, químico o psicológico. Es decir, no se debe confundir la confiabilidad del instrumento de medida con la estabilidad o modificabilidad del constructo medido, por lo que el usuario de los test tendrá que identificar las fuentes de error que afectan a las mediciones y no achacar por ejemplo, la baja confiabilidad de un instrumento de medida a lo que puede ser sencillamente variabilidad legítima de la variable de medida. Lo que debe quedar claro es que la confiabilidad se refiere a la estabilidad de las mediciones cuando no existen razones teóricas, ni empíricas para suponer que la variable a medir haya sido modificada diferencialmente para

los sujetos, por lo que se asume su estabilidad, mientras no se demuestre lo contrario (Muñiz, 1992, p. 31-32)

#### **1.4.1.3 Métodos de la TCT para estimar la confiabilidad**

Dentro de la TCT, la confiabilidad viene expresada por coeficiente de confiabilidad, que no es más que el coeficiente de correlación momento-producto de Pearson de las puntuaciones de un test consigo mismas. Dentro de ésta se proponen tres métodos para estimar la confiabilidad de los test; son las siguientes:

1. Método test-retest. Implica la aplicación de un mismo test en dos ocasiones diferentes al mismo grupo de sujetos. Las correlaciones entre las puntuaciones obtenidas de las dos aplicaciones, será el coeficiente de confiabilidad. Debido a que este método es congruente con el modelo, se le denomina a la estimación obtenida coeficiente de estabilidad, pues indica en qué grado son estables las mediciones realizadas en la primera aplicación del test (Muñiz, 1992; García, 1993).
2. Método de las formas paralelas o Coeficiente de confiabilidad. Se define como la correlación entre las puntuaciones obtenidas por los sujetos en dos formas paralelas de un test, esto significa que al construir un test que se aplicará a un grupo normativo pertinente se construya otro que mida lo mismo que el anterior, con el mismo formato, el mismo número de ítems y del mismo tipo, el cual se aplica al mismo grupo y la confiabilidad vendrá expresada por el coeficiente de correlación de Pearson entre las puntuaciones obtenidas de

ambos test (García, 1993) Estas dos formas paralelas son un indicador de la estabilidad de las medidas: siempre y cuando si no hubiese errores aleatorios de la medida, la correlación debería ser perfecta; es decir  $P_x x' = 1$ , por lo tanto el grado en que  $P_x x'$  se aleja de 1 nos indicará que medida nuestras mediciones están afectadas por errores aleatorios de medida, siempre en el supuesto de que las dos formas  $x$  y  $x'$  realmente lo sean (Muñiz, 1992, p.32).

3. El método de las dos mitades. Se considera el más sencillo de los tres para estimar el coeficiente de confiabilidad. Por este método se aplica el test una sola vez, obteniéndose para cada sujeto las puntuaciones correspondientes a cada una de las dos mitades en las que se divide el test. El coeficiente de confiabilidad viene dado por la correlación entre esas dos mitades (que sería la estimación de la confiabilidad del test mitad) más una corrección para obtener la confiabilidad del test total. La estimación así obtenida, más que equivalencia o estabilidad, como en los casos anteriores indica covariación o consistencia interna del test (Muñiz, 1992, p.34).

#### **1.4.1.4 Ventajas y desventajas de los métodos**

Como se pudo observar la lógica de estos tres métodos es clara, sin embargo se plantean diversos problemas experimentales referentes a la validez interna, la cual en última instancia queda a criterio del psicólogo o del usuario de los test para cada situación que se le presente, por ello es importante que se consideren las ventajas y desventajas que cada uno de estos métodos presenta:

En el método test-retest su principal ventaja es que es un procedimiento poco costoso, sencillo y fácil de llevar a la práctica. además es adecuado para cualquier tipo de test; sus inconvenientes es el delimitar en primera instancia el tiempo óptimo que debe transcurrir entre ambas aplicaciones "si se deja mucho tiempo se puede introducir una gran fuente de invalidez interna, a saber, la ignota influencia diferencial de ese periodo de tiempo en los sujetos: pero si transcurre poco tiempo, la invalidez interna se cuele vía memoria de lo realizado previamente" (Muñiz, 1992, p. 34).

Otros aspectos que puede influir en los resultados es el problema de la fatiga si el periodo de tiempo entre una aplicación y otra es muy corto o inexistente; la pérdida de los sujetos del grupo normativo sería otro problema, al dejar pasar mucho tiempo entre la primera y segunda aplicación, lo que generaría la mortandad experimental, y la variabilidad del rasgo (aspecto de la conducta que se está midiendo) que puede sufrir modificaciones con el paso del tiempo, con lo que una correlación poco elevada no sólo podría interpretarse como debida a un test poco fiable, sino también a una variación en el rasgo que se está midiendo (García, 1993, p.26)

En el método de formas paralelas, el problema fundamental se presenta en la construcción de dichas formas paralelas, ya que resulta difícil a nivel teórico hacer un test que mida exactamente lo mismo que el otro, además que en la práctica resulta sumamente laborioso, sin embargo se puede considerar que este método sería el más recomendable para determinar el grado de confiabilidad del test

En cuanto al método de dos mitades se puede decir que es el más funcional de los tres, ya que exige una sola aplicación; elimina el problema del cansancio, el aprendizaje y la mortandad experimental. Su principal desventaja es que sólo es aplicable a los test de potencia y a los libres de tiempo.

Finalmente para concluir este apartado se considera conveniente mencionar algunas limitaciones que tienen las TCT.

#### **1.4.1.5 Limitaciones de la teoría clásica de los test**

- a) Si es cierto que la TCT ofrece un coeficiente de confiabilidad global para todas las pruebas, los datos empíricos acumulados muestran que la precisión de las pruebas está en función del nivel del sujeto en la variable medida. Es decir, los test no miden con la misma precisión a todos los sujetos. En otros términos la teoría clásica, no puede predecir como responderá un individuo a los ítems a menos que éstos se hayan aplicado con anterioridad a personas similares (Díaz, 1996).
- b) Las mediciones no son invariantes respecto del instrumento utilizado. Por ejemplo si se utilizan tres tests de inteligencia distintos para evaluar a tres sujetos, en paridad no se puede comparar el nivel de inteligencia de los tres sujetos evaluados.
- c) En el marco de la TCT, las propiedades de los instrumentos de medida no son invariantes respecto de los sujetos utilizados para estimarlas. Así por ejemplo la dificultad de los ítems está en función de las personas utilizadas para calibrarlos (Muñiz, 2001).

A pesar de dichas limitaciones el modelo lineal clásico sobre el que se asentó la TCT, sigue siendo un modelo que da cobertura hasta nuestros días a las necesidades de los profesionales de la medición psicológica, tanto en lo relativo a la confiabilidad de las mediciones (estimación de error) como a la validez. Pero también es importante decir, que estas limitaciones nos muestran las fuentes de error o variabilidad que se presenta en los test construidos bajo la TCT y que tratarán de ser resueltas por la Teoría de la Generabilidad y la Teoría de la Respuesta al Ítem.

#### **1.4.2 Teoría de la Generabilidad (TG)**

Tiene sus antecedentes teóricos en la psicometría, la cual se encuentra descontenta con la TCT en cuanto a los resultados que ésta había tenido hasta el momento, por lo que la TG constituye un intento de ampliar el acercamiento clásico a la confiabilidad, aplicando para ello otras técnicas del análisis estadístico. "La TG pone el acento en la estimación de los diversos *componentes de la varianza*, más que en las pruebas de *significación estadística o contrastes de hipótesis*" (Martínez, 1995, p. 171).

En la TG una medida es considerada como una muestra independiente proveniente de un universo de medida y cada medida de este universo se describe en función del número de condiciones bajo las cuales ha sido obtenida la medida. Los diferentes tipos de condiciones a diferentes componentes de varianza (García, 1993, p.70).

Este aspecto es lo que le da una ventaja clara a la TG frente a la TCT, ya que permite estimar la confiabilidad de un instrumento de medida en situaciones en las que intervienen múltiples fuentes de error o variabilidad de las puntuaciones. Lo que en la TCT está determinado en función de las diferencias individuales y los errores aleatorios, tendiendo a uniformar las condiciones de medida (en el mismo instrumento o en sus formas paralelas) lo que genera dos problemas: "El primero es que en la práctica, nunca nos encontramos con condiciones de medida estrictamente paralelas: las llamadas formas paralelas no siempre presenta una equivalencia respecto a contenido, media y varianza, y, por otra parte, resulta del todo imposible uniformar las situaciones de aplicación. El segundo problema es que, en el caso utópico de lograr condiciones de medida idénticas, estaríamos sesgando las puntuaciones de los sujetos, es decir, dichas condiciones podrían beneficiar sistemáticamente a algunos sujetos y perjudicar a otros" (Muñiz, 1992, p. 76). Por tanto lo que va a aportar la TG para superar estos problemas es la utilización de una serie de técnicas para estudiar el grado en que una serie de medidas realizadas sobre un determinado grupo pueden ser generalizadas y hechas extensivas a un grupo de sujetos diferentes al primitivo.

Para lograr lo anterior la TG introdujo una serie de modificaciones importantes a la TCT, siendo las más significativas las siguientes:

- 1) Utiliza el concepto estadístico de muestreo de fuentes de variación múltiples, lo que permitirá tratar cada característica de la situación de medición como una faceta de un plan de medición concebido como un diseño experimental. Cada una de estas facetas tiene su propia variabilidad, lo que permitirá considerar distintas fuentes de error, que pueden ser tratadas utilizando los procedimientos del Modelo Lineal General del Diseño Experimental.
- 2) Sustituye el concepto de medidas paralelas por el de medidas aleatoriamente paralelas, considerando que las distintas condiciones o

componentes de un procedimiento de evaluación,(...) pueden considerarse una muestra aleatoria de un universo más amplio, definido por las condiciones de muestreo bajo las que se obtiene cada medición concreta. Las concepciones de medidas aleatoriamente paralelas es menos restrictivo que la noción de paralelismo de la TCT

- 3) Sustituye el concepto de puntuación verdadera por el de puntuación de universo, definida como la esperanza matemática de todas las mediciones posibles y el error como la fluctuación muestral correspondiente a la extracción aleatoria de ciertos niveles de las características consideradas (calificadores, momentos, conjuntos de ítems particulares, situaciones, etc.)
- 4) Para definir la confiabilidad o grado en que una puntuación se aproxima a la puntuación del universo habrá que determinar su grado de generalidad o invarianza dentro de la población de condiciones. Se sustituye el concepto de confiabilidad por el de generalizabilidad del resultado concreto a la población de condiciones (Martínez, 1995, p. 170).

#### **1.4.2.1 Marco Conceptual.**

La teoría de la Generalizabilidad asume que todas sus condiciones de un test, están aleatoriamente extraídas de un enorme conjunto o universo de condiciones, es decir, en la construcción de un test, la selección de los ítems, se realiza de una muestra aleatoria de los ítems que componen el universo general de todos de los ítems posibles que puedan medir el rasgo que se desea evaluar (García, 1993). Por tanto el universo de generalización es la población de medidas que pueden ser obtenidas midiendo bajo todas las condiciones posibles. Por ejemplo supónganse que se desea diseñar un procedimiento para medir la inteligencia general de una determinada población, la medida se realizará bajo diferentes condiciones (distintos ítems, distintos examinados, etc.) y seguramente las puntuaciones de cada sujeto variarán de una condición a otra.

Para explicar dichas variaciones, a TG considera que existen por lo menos dos factores o facetas<sup>2</sup> que producen dicha variabilidad en los resultados: uno son los ítems y el otro los examinados. Para diferenciar a los sujetos el investigador tendría que considerar estos dos factores como fuentes de error. En donde habría:

- 1) faceta de diferenciación que serían los sujetos.
- 2) Faceta de generalización que la constituirían los ítems y examinados y el universo de observación admisible que estaría constituida por todos los ítems posibles y bajo todas las condiciones de medida aceptables (Muñiz, 1992, p. 76).

Es decir, los resultados obtenidos con un test de inteligencia construido con un conjunto aleatorio de ítems extraídos del universo formado por todos los posibles ítems que midan la inteligencia, pueden generalizarse a todos los test de inteligencia construidas con ítems que sean muestras aleatorias y representativas del mismo universo

Dentro de la TG se pueden distinguir dos tipos de estudio:

---

<sup>2</sup> El término faceta fue introducido por Gronbach, para designar a cada característica de la situación de medida que pueden modificarse de una medición a otra y que en consecuencia puedan hacer variar el resultado obtenido. Ejemplo de ello sería elementos del test; observadores, evaluadores, formas de codificación, etc (Martínez 1995, p. 171)

#### 1.4.2.2 Estudios de generalizabilidad (estudio G)

Van dirigidos a generalizar una muestra al universo de medidas: la estabilidad de las respuestas a través del tiempo, la equivalencia de las puntuaciones entre dos o más formas de un mismo instrumento de medida, las interacciones de los ítems de un test, son ejemplos de estudios tipo G (García, 1993, p. 72).

A partir de los datos obtenidos del estudio G se puede plantear un estudio de decisión ( estudio D), cuyo objetivo fundamental es la optimización de las condiciones de medida. En función de los componentes de la varianza obtenida en el estudio G se puede predecir o aumentar el número de niveles de una faceta. "Si la varianza atribuible a una de generalización es grande, esa faceta contribuye al error de una manera significativa, por lo que se requerirá aumentar el número de niveles de la misma si se desea un nivel alto de generalizabilidad. Si por el contrario es pequeña, el investigador puede reducir el número de niveles de esa faceta sin que por ello se reduzca la precisión en la generalización, o se puede incluso decidir fijarla, con lo que no produciría variabilidad en los resultados y por lo tanto no podría ser utilizada como faceta de generalización" (Muñiz, 1992, p 77)

La diferencia entre estos dos tipos de estudio es que un estudio G se realiza con el fin de analizar la constancia de medida de una determinada variable a través de distintas condiciones, mientras que en los estudios D su finalidad es decidir sobre el procedimiento idóneo de medida de dicha variable o tomar alguna decisión sobre los sujetos examinados.

La técnica que va a permitir que se alcancen los objetivos de cualquiera de los estudios mencionados es la de *análisis de varianza* que va a permitir estudiar

de forma adecuada las distintas fuentes de variación y estimar el coeficiente de confiabilidad o generabilidad, (procedimiento más adecuado que el coeficiente de confiabilidad de la TCT) que es el cociente entre la varianza de las puntuaciones universo y las de las puntuaciones empíricas. Lo que fundamentalmente le preocupa a la TG es : cómo estimar mejor la puntuación universo de un sujeto y cuál es el error que se comete en esa estimación. Al respecto Muñiz (1992) plantea que se distinguen tres clases de error al partir del supuesto de que los test paralelos pueden tener distinta medida por lo que la medida de puntuaciones empíricas no tiene porque coincidir con la medida de puntuaciones universo. Estas tres clases de error son:

“ $\Delta$ : Diferencia entre la puntuación empírica de un sujeto y su puntuación diferencial universo. Incluye el error del test (la diferencia entre la medida del test y la media de las puntuaciones universo)

$\delta$ : Diferencia entre la puntuación diferencial empírica de un sujeto y su puntuación diferencial universo. Al utilizar puntuaciones diferenciales se elimina el error del test, es decir, a diferencia entre la medida del test y la medida de puntuación universo.

$\epsilon$ : Error de estimación cometido al pronosticar la media de las puntuaciones universo a partir de la media de las puntuaciones empíricas” (p.78).

Por la magnitud de las distintas fuentes de variación y las múltiples fuentes de errores en la TG se cuenta con un modelo estadístico que permite llevar a cabo esta diferenciación y es el modelo de análisis de varianza (ANOVA) dentro del cual existen dos modelos básicos para llevar a cabo una partición de la variación total de las puntuaciones en fuentes independientes de variabilidad y son:

1) Diseños de una sola faceta; dentro de éste, se considera como faceta de diferenciación a los sujetos y como faceta de generalización los ítems. Existen dos tipos de diseño de una sola faceta:

- *Diseño cruzado*  $s \times i$  –Consiste en aplicar la misma muestra aleatoria de ítems a todos los sujetos. (Si determinamos  $N_s$  al número de sujetos y  $N_i$  al número de ítems, daría como resultado  $N_s \times N_i$ , observaciones)
- *Diseño anidado*  $i : s$ . Consiste en aplicar a cada sujeto una muestra aleatoria diferente de ítems. Como los ítems se reparten entre los sujetos el número de observaciones será igual a  $N_i$ .

2) Diseños de dos facetas, dentro de éstos, se consideran las dos facetas anteriores (diferenciación de sujetos y generalización de los ítems del test) más los examinadores. Parte también de diseños cruzados, en el que se agrega el tercer componente a la fórmula:  $s \times i \times j$ , que permite obtener los estimadores de las varianzas para cualquier tipo de diseño, utilizando la suma de cuadrados para obtener estimadores de los componentes de varianza (Muñiz, 1992).

Cada uno de estos diseños presenta sus respectivos procedimientos estadísticos para obtener la viabilidad y el grado de error de los test utilizados en investigación, dado que no es el interés de este trabajo desarrollar dichos procedimientos, se concluye este apartado estableciendo las ventajas que presenta la TG frente a la TCT. Una de esas ventajas es que en la TG para su aplicación exige el cumplimiento de menos supuestos estadísticos. El único supuesto a cumplir por ésta, es el muestreo aleatorio de sujetos y condiciones de

medida La diferencia más clara entre ambas teorías la encontramos en el concepto de error. La TCT trata el error que afecta las puntuaciones de los sujetos en un test como un constructo unitario. Así, la confiabilidad test-retest sólo considera sólo considera como fuente de error el momento de aplicación, los coeficientes de consistencia interna sólo consideran el muestreo de los ítems y el coeficiente de confiabilidad obtenido por el método de las formas paralelas considera ambos factores, pero sin diferenciarlos. La TG, considera por separado cada una de las fuentes de error, así como sus interacciones. Partiendo de un estudio G: s x i x o (sujeto x ítems x ocasiones), la TG ofrece la posibilidad de calcular los tres tipos de coeficientes: estabilidad (test-retest), constancia (alfa) y equivalencia (formas paralelas)" (Muñiz, 1992, p.98).

#### **1.4.3 Teoría de la Respuesta al Ítem. (TRI)**

Esta teoría se desarrolló para resolver varios problemas que presentaba la TCT y que no habían sido resueltos de forma satisfactoria. Algunos de esos problemas son:

1. "El uso de índices de los reactivos cuyos valores dependen de la población particular de la cual se obtuvieron.
2. La estimación de la habilidad del examinado depende del conjunto específico de reactivos incluidos en la prueba" (Díaz, 1992).

Es decir, que los resultados de una medición realizada sobre una variable depende por un lado del test utilizado y por otro del grupo normativo a partir del cual se estimaran los distintos parámetros de ítems, en términos más concretos, es como tratar de obtener el peso de una persona en función de un tipo de báscula

específica, y además consideran que las cualidades y características de la balanza se establecerán en función del sujeto al que se pesa.

Desde el punto de vista de la TRI estos problemas, no se han podido resolver bajo los supuestos de la TCT debido a que ésta presenta las siguientes limitaciones:

- Dependencia de un grupo normativo, teniendo como consecuencia que el valor de los estadísticos tales como el poder discriminativo de un ítem o su índice de dificultad depende de la habilidad media del grupo en que sean calculados.
- Los valores manejados dentro de la TCT como coeficiente de confiabilidad, coeficiente de validez, índices de homogeneidad y validez de los elementos, se estiman mediante el coeficiente de Pearson, lo que hace que dichos valores estén directamente relacionados con la variabilidad de las puntuaciones del test en el grupo normativo. Es decir existen tantos coeficientes diferentes (de confiabilidad, validez, etc ) como aplicaciones del test sean hechas a distintos grupos
- Un tercer inconveniente surge cuando se quiere establecer comparaciones entre las aptitudes de los sujetos que hayan sido medidos en la misma variable por test diferentes.
- Uno de los conceptos fundamentales de la TCT es el de confiabilidad, que se define en términos de correlación entre formas paralelas de un mismo test, lo que resulta empíricamente muy difícil de conseguir.
- La estimación de la verdadera puntuación de un sujeto en el test, los intervalos confidenciales en torno a la puntuación empírica obtenida por un sujeto en un test y otras medias y comparaciones se basan, fundamentalmente, en el error típico de medida del test. Este error típico se supone constante para un mismo test para cualquier muestra de sujetos. Este supuesto resulta muy débil y muy difícil de mantener, con lo que todo el sistema de puntuaciones verdaderas y de comparaciones entre puntuaciones de los tests según la TCT, falla desde la misma base (García, 1993, p.142-143)

La solución a estos problemas citados y otros más, va a venir de la mano de nuevos modelos psicométricos, denominados genéricamente Teoría del Rasgo a los Items (TRI), los cuales se inician en los años 50's, se incuban en los 70's y florecen en los 80's y los 90's, dando un giro radical a la forma de hacer psicometría. Veamos el proceso histórico que sigue la TRI.

#### 1.4.3.1 Antecedentes Históricos

Se considera que surge en los años 50's y 60's como reacción a la TCT justificada por las diferencias que presentaba y que podían ser resueltas a través de la nueva teoría; su origen se puede establecer en el trabajo publicado por D.N Lawley en 1943, el cual se titula "*On problem connected with item selection and test construction*" en el que se presenta una primera descripción del modelo de ojiva normal, como función que pone en relación un rasgo latente tomado como continuo en la probabilidad de respuesta correcta a un ítem (Melía, Tomas y Oliver p. 108)

Después de la presentación de este trabajo aparecen diversas publicaciones que toman como punto de partida los modelos del rasgo latente. Una de las más importantes y que se considera de gran influencia para la psicometría, es el libro publicado por Soufler, Guttman, Suchma, Lazarsfeld y Sten, "*Measurement and prediction*", que aparece en 1950 y en donde sobre sale la aportación de Lazarsfeld, quien elabora uno de los primeros trabajos sobre la teoría de la estructura latente, este autor introduce en término "rasgo latente" y el concepto de "curva característica del ítem", que se convierte en el concepto central de la teoría y alrededor del cual giran todos los desarrollos posteriores (por lo menos durante los 30 años siguientes). "Esta curva relaciona la probabilidad de una respuesta a un ítem habitualmente dicotómico, condiona a nivel del rasgo latente que posee el sujeto". (Melía, Tomas y Oliver, 1992, p.108)

En 1957, Birnbaum presenta una nueva función matemática para describir la función característica del ítem, la función logística, derivada del campo del bioensayo. Esta función logística difiere levemente de la ojiva normal, en términos de que posee una mejor tratabilidad matemática, por lo que tendió a sustituirla en la investigación. Al igual que el modelo de la ojiva normal, la función logística queda determinada por tres parámetros: "la dificultad del ítem, la capacidad de discriminación del ítem y el parámetro de conjetura que da cuenta de la probabilidad de responder por azar". (Meliá, 1992, p. 109) Este modelo será conocido también como modelo logístico o de ojiva de tres parámetros.

Hacia los años 60's casi de manera paralela a los trabajos desarrollados en Estados Unidos con Lord y Birnbaum, el danés Rach, desarrolla un modelo (el cual se conoce con el nombre de modelo logístico de Rach) de un solo parámetro, que también tuvo gran influencia en el desarrollo de diversos trabajos tanto teóricos como prácticos en el campo de la psicometría, lo característico del modelo es que la respuesta del sujeto sólo se ve influida por dos parámetros, uno del ítem que se le presenta y el otro de la capacidad del sujeto, se podría considerar que este modelo queda incluido en el tercer parámetro del modelo antes mencionado, en el que se asume que no hay probabilidad de contestar al azar (conjetura igual a 0) y todos los ítems discriminan lo mismo (parámetro de discriminación igual a constante), aspectos que hasta la fecha siguen generando fuertes polémicas en la literatura.

Sobre la misma línea de los modelos logísticos en 1967 McDonal, propone un modelo de cuatro parámetros, el cual pretende explicar los casos en que

los individuos de alta capacidad no responden correctamente a ítems fáciles. Este parámetro representará la probabilidad de contestar correctamente el ítem para sujetos con capacidad muy alta se considera que el desarrollo de este modelo responde a un interés puramente teórico.

Para la década de los 70's los modelos desarrollados hasta el momento habían llegado a su consolidación teórica, por lo que el énfasis de los trabajos desarrollados en esta década son de índole más aplicada, centrados sobre todo en los problemas concretos de implementación del modelo, estimación y ajuste a los datos. Este interés tiene su sentido de ser, debido a que para estos años empiezan a aparecer programas informáticos, sin los cuales la aplicación práctica real de éstos, era casi imposible. Uno de los programas estadísticos más influyentes para el manejo de los cálculos que se requieren para la aplicación de los modelos antes descritos es el LOGIST, que permite la estimación con procedimientos de máxima verosimilitud conjunta la capacidad y parámetros del ítem para modelos de tres parámetros; de uno y dos parámetros con restricción para éstos dos últimos.

En las décadas siguientes (80's y 90's) la disponibilidad de estos programas llevó a su implementación en grandes muestras, lo que genera una gran producción de investigaciones y la aplicación en situaciones más concretas de la TRI, especialmente en campos en los que parece que sus resultados son superiores a los obtenidos mediante la TCT, siendo algunos de éstos; igualación vertical y horizontal, bancos de ítems y estudios de sesgos ( Melia, Tomas y Oliver, 1992, p. 114).

Este recorrido general sobre los antecedentes históricos de la TRI, nos remite a la delimitación de su definición:

### **1.4.3.3 Definición**

El nombre de TRI, proviene de que este enfoque se basa en las propiedades de los ítems más que en las del test global, conocida también como teoría del rasgo latente, teoría de la curva característica del ítem, lo cierto es que la denominación universal es la de TRI y ello, efectivamente, porque refleja el funcionamiento real del modelo basado en los ítems, permitiendo además distinguirlo de otros acercamientos más generales que utilizan el concepto de rasgo latente en psicología, como puede ser el de análisis factorial, análisis multidimensional o ecuaciones estructurales (Hambletan y Swanathan, 1985 en Muñiz, 1990, p.13)

En términos más concretos se podría decir que la TRI "supone que una conducta individual puede ser explicada, en buen grado, definiendo ciertas características humanas (no observables) denominadas rasgos, estimando 'cuantitativamente' la posición del individuo en cada uno de estos rasgos y utilizando los valores numéricos obtenidos de dicha estimación para predecir o explicar la conducta individual en situaciones pertinentes. Es decir, se trata de especificar las relaciones entre las puntuaciones observables obtenidas por un sujeto en un test y 'los rasgos latentes' que se supone han de subyacer para poder explicar dichas puntuaciones, en concreto, la habilidad que dicho individuo ha de poseer, para poder explicar sus respuestas a los ítems del

test (García, 1993, p. 143) Por tanto sus principales objetivos están centrados en:

- 1) Obtener mediciones que no varíen en función del instrumento utilizado. que sean invariantes respecto de los test empleados
- 2) Disponer de instrumentos de medida cuyas propiedades no dependan de los objetivos medidos, sean invariables respecto de los sujetos evaluados (Muñiz, 1990, p 16).
- 3) Estimar los parámetros, incluso los errores estándar.
- 4) Elegir el mejor valor de  $\theta$  para un patrón determinado. Mejor es definido con frecuencia, pero no de manera universal, como una estimación de probabilidad máxima (Nunnally, 1995, p. 445).

Los defensores de esta teoría consideran que es de gran importancia por el alcance que tiene para la comprensión de los instrumentos de medición (test) y sobre todo para definir con mayor precisión su grado de error.

De acuerdo con Nunnally (1995) son cuatro las razones para estimar la importancia de la TRI:

- a) Una TRI puede comparar dos pruebas que constan explícitamente de reactivos diferentes. En consecuencia permite hacer comparaciones entre diferentes ocasiones para el mismo sujeto donde el recuerdo de repuestas previas es un problema, aún así las dos pruebas no tienen reactivos en común. Esto se denomina medición de prueba libre y es importante para la evaluación hecha a la medida y la evaluación adaptativa computarizada
- b) Puede mostrarse que sujetos con el mismo puntaje clásico (número de correctas) difieren en  $\theta$ , según los supuestos hechos por el modelo TRI.
- c) La estimación clásica del nivel del atributo (habilidad) no se relaciona linealmente con  $\theta$ .
- d) Las estimaciones clásicas de dificultad y discriminación tales como probabilidad de una respuesta correcta, valor  $p$ , y la correlación reactivo-total ( $r_n$ ) no son independientes entre sí, ya que dependen de las

habilidades de los sujetos. Por tanto, un reactivo cuyo valor  $p$  es 5 en una población general, tendrá un valor inferior entre los menos hábiles, pero la psicometría clásica no puede predecir la magnitud de la disminución. En principio, las estimaciones TRI correspondientes no resultan afectadas con estos problemas. (p.445)

Los elementos anteriormente expuestos permiten comprender con mayor claridad porque la TCT no permite hacer predicciones acerca de cómo se comportará un individuo o un grupo particular ante un reactivo dado, lo que sí es factible de definir a través de la TRI, de ahí de importancia ya que la probabilidad de predicción que se puede obtener es más precisa para una gran variedad de situaciones como por ejemplo, el comportamiento de un profesional ante distintas tareas.

Según Hambleton (en Díaz, 1996) son dos los principios básicos que sustentan a la TRI y que permiten hacer dichas predicciones:

- 1) El desempeño de un examinado en una prueba pueden predecirse por un conjunto de rasgos, rasgos latentes y habilidades; y
- 2) La relación entre las respuestas de los examinados a los reactivos y el conjunto de rasgos que subyacen a la respuesta ante el reactivo, pueden describirse por una función monotónicamente incrementada llamada función específica que, en la medida en que el nivel del rasgo se incrementa, también lo hace la probabilidad de una respuesta correcta ante el reactivo (p.57).

Los modelos matemáticos descritos permiten plantear la probabilidad de que un sujeto responda de modo correcto a un reactivo, considerando por un lado la habilidad del sujeto y por otro las características del reactivo, de estos dos

elementos va a depender dicha predicción sustentándola bajo los supuestos siguientes:

- Sólo puede medirse una habilidad por el reactivo que forma parte del examen; este es el supuesto de unidimensionalidad.
- La función característica del reactivo refleja la relación real que hay entre las variables no observables (habilidad) y las observables (las respuestas a los reactivos) (Díaz, 1996, p.57)

Veamos con más detenimiento el fundamento de estos supuestos.

#### **1.4.3.3 Curva característica de los ítems**

Los modelos del TRI asumen que existe una relación funcional entre los valores de la variable que miden los ítems y la probabilidad de acertar a éstos, denominando a dicha función curva característica de los ítems (Muñiz, 1990). La cual se puede definir "como una función matemática que relaciona la probabilidad de contestar correctamente un ítem con la habilidad medida por el conjunto de ítems que configuren el test, en otros términos es la función de regresión no lineal de la puntuación en el ítem sobre el rasgo o habilidad medida por el test" (García, 1993, p.146-147); lo que significa que la probabilidad de acertar a un ítem sólo depende de los valores de la variable medida por el ítem; por lo tanto, sujetos con distintas puntuación en dicha variable tendrán probabilidades distintas de superar determinado ítem. La curva característica de los ítem permanecerá invariable de un grupo de sujetos a otro, siempre que el espacio el espacio latente de ambos

grupos sea el mismo y esté completo, es decir, de determinan a partir de los siguientes parámetros:

#### 4.3.3.4 Parámetros de la curva característica del ítem

Las respuestas dadas por cualquier sujeto al test se explican por la posesión de un habilidad única y unidimensional (la cual se explicará más adelante), y los times que se estudien bajo los supuestos de la TRI serán dicotómicos o estarán dicotomizados. Para poder definir adecuadamente lo anterior se deben tener en cuenta los tres parámetros siguientes:

*Parámetro a.* Se denomina índice de discriminación y su valor es proporcional a la pendiente de la recta tangente a la CCI en el punto de máxima pendiente de ésta, cuanto mayor sea la pendiente, mayor será el índice de discriminación

*Parámetro b.* Se denomina índice de dificultad y es el valor correspondiente al punto de máxima pendiente de la CCI.

*Parámetro c.* Representa la probabilidad de acertar el ítem al azar cuando 'no se sabe nada', es decir, el valor de  $p(\theta)$  cuando  $\theta = -\infty$ . En otras palabras, es el valor asintótico de la CCI cuando  $\theta$  tiende a  $-\infty$ .

La CCI queda definida cuando se especifican estos tres parámetros y se adopta una determinada función matemática para la curva" (Muñiz, 1990, p.20-22).

### Unidimensionalidad e independencia local

Se asume que las respuestas dadas en un test por un sujeto se explica por las habilidades específicas o rasgos latentes. También se señaló la relación funcional que establece la Cci entre la probabilidad de acertar aun ítem y los valores  $\theta$ . "Por tanto, si el modelo es correcto la probabilidad de acertar a un ítem únicamente

dependerá de un factor, a saber de  $\theta$ . En otras palabras, la TRI asume implícitamente en su formulación que los ítems destinados a medir la variable  $\theta$  constituyen una sola dimensión, "unidimensionales" (Muñiz 1990, p.23)

¿ Pero cómo se comprueba la unidimensionalidad?; tan sólo Hattie (en García, 1993) distingue 87 índices para su estudio, sin embargo, la técnica que parece más adecuada para probar dicho supuesto es el análisis factorial, aunque tropieza con la dificultad de tener que someter ítems dicotómicos, lo que se resuelve realizando dicho análisis sobre la matriz de correlaciones  $\Phi$  o de correlaciones tetracóricas. Se dice que la unidimensionalidad sería perfecta si el primer factor extraído explicará el 100% de la varianza. Como esto no es posible la unidimensionalidad se convierte en un problema de grado, es decir, cuanto mayor sea la proporción de la varianza, explicada por el primer factor con respecto a los demás, menos riesgo se correrá de violar el supuesto de la unidimensionalidad, y si se cumple se dará la *independencia local* entre los ítems (García, 1993, p.145-146); esto es, que para un sujeto con un determinado valor en la variable unidimensional su respuesta a un ítem no viene influida por sus respuestas en los otros, por tanto, la independencia local puede expresarse diciendo que la probabilidad de que un sujeto acierte un ítem es igual al producto de las probabilidades de acertar a cada uno de ellos. Análogo a lo anterior, se puede decir también que hay independencia local de los sujetos en el sentido de que el rendimiento de un sujeto es independiente del rendimiento de los otros (Muñiz, 1990, p.25)

Como se ha visto hasta el momento la CCI es el rasgo esencial de la TRI, y para su explicación se adoptan una u otra función matemática según se tenga en

cuenta los parámetros de los ítems antes descritos, estos parámetros deberán de responder a diferentes modelos los cuales se han centrado en dos tipos de funciones matemáticas para la CCI y son: la función logística y la curva normal acumulada, lo que da lugar a seis modelos, según se cumplan o no dos de los tres parámetros para cada una de las dos funciones.

Debido a la mayor estabilidad matemática que tiene la función logística, son éstos en los que se centrará la atención debido a que se dispone de más datos.

#### **1.4.3.5 Modelos**

##### *Modelo logístico de un parámetro. (Modelo de Rasch)*

Formulado por Rasch en 1960, se considera como el más popular de la TRI, debido a la sencillez emanada de su lógica: la respuesta a un ítem sólo depende de la competencia del sujeto y de la dificultad del ítem es decir, de  $\theta$  y de  $b$ .

Según dicho modelo, la CCI viene dada por la función logística, y el único parámetro de los ítems a tener en cuenta es  $b$ , el índice de dificultad. En términos matemáticos el modelo quedaría representado en la siguiente expresión; para el caso de un parámetro:

$$P_i(\theta) = \frac{e^{D(\theta - b_i)}}{1 + e^{D(\theta - b_i)}}$$

Donde:  $P_i$ : Probabilidad de acertar el ítem  $i$  a determinado nivel  $\theta$ .

$\theta$ : Valores de la variable medida

$b_i$ : índice de dificultad del ítem  $i$ .

$e$ : Base de los logaritmos neperianos (2,72)

$D$ : constante. Cuando toma el valor 1,7, la función logística se aproxima a la fórmula acumulada.

#### *Modelo de dos parámetros.*

Desarrollado por Birnbaum (1968) Asume que la CCI viene dada por la función logística y tiene en cuenta dos parámetros: dificultad del ítem ( $b$ ) y su capacidad para discriminar ( $a$ ). Su fórmula viene dada por

$$P_i(\theta) = \frac{e^{D a_i (\theta - b_i)}}{1 + e^{D a_i (\theta - b_i)}}$$

Donde:  $P_i$ : Probabilidad de acertar el ítem  $i$  para valor

$\theta$ : Valores de la variable medida

$b_i$ : índice de dificultad del ítem  $i$

$a_i$ : índice de discriminación del ítem  $i$

$e$ : Base de los logaritmos neperianos (2,72).

D: constante. Cuando toma el valor 1.7, la función logística se aproxima a la fórmula acumulada.

### *Modelo de tres parámetros.*

Se obtiene del dos con la adición del tercer parámetro, comúnmente denominado  $c$ , relativo a la probabilidad de acertar el ítem al azar, cuando no se conoce la respuesta. Técnicamente  $c_i$  es el valor de  $P_i(\theta)$  para un valor de  $(\theta) = -\infty$ . El modelo se expresa así:

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{\frac{D a_i (\theta - b_i)}{1 + e^{D a_i (\theta - b_i)}}}}{1 + e^{\frac{D a_i (\theta - b_i)}{1 + e^{D a_i (\theta - b_i)}}}}$$

Donde  $P_i(\theta)$ ,  $e$ ,  $D$ ,  $a_i$ ,  $\theta$  y  $b_i$  tienen la misma significación que la ya citada para el caso de dos parámetros, y  $c_i$  es el valor ya citado de  $P_i(\theta)$ , cuando  $\theta = -\infty$ .

Con la intención de no dejar a un lado los modelos de ojiva normal, se plantea que la CCI viene dada por la función de la curva normal acumulada, si bien son modelos que preceden a los modelos logísticos, éstos son más utilizados, como ya se dijo por su mayor manejabilidad matemática.

Finalmente y para cerrar este capítulo, se mencionan algunas de las limitaciones que presenta la TRI. Debido al uso de criterios estadísticos para la selección de los reactivos no se asegura una prueba con contenido válido por completo. Las diferencias en los procedimientos de selección de los contenidos

pueden generar una prueba con un bajo nivel de validez de contenido. Otro problema es que cuando se utilizan funciones de información de los reactivos durante el desarrollo de una prueba, es probable que los valores se sobrevaloren y por tanto la función de la información podría sesgarse (Díaz, 1996)

Por otra parte el problema de la unidimensionalidad tanto en la TCT y la TRI es una seria objeción a estos modelos, ya que muchas conductas humanas son claramente complejas y difíciles de analizar en todas sus dimensiones primarias aislando una de otras. También en muchas ocasiones la variable tiempo resulta ser incompatible desde el punto de partida con el supuesto de unidimensionalidad, ya que el tiempo es por sí mismo una clara dimensión (García, 1993)

Lo que es innegable es que la psicometría es un instrumento más en las manos del profesional de la psicología, que puede ayudarle a ser más objetivo y cometer menos errores en la toma de decisiones. Considerando la importancia que ésta tiene, ahora nos centraremos en uno de los fenómenos más interesantes y que ha sido blanco favorito de la psicometría: la inteligencia y su medición

## CAPÍTULO 2

### INTELIGENCIA Y SU MEDICIÓN

Uno de los fenómenos psicológicos que tantas dificultades ha presentado tanto en su definición como en su comprensión es el de la inteligencia. Esta característica es lo que posiblemente ha impulsado también a que se haya estudiado y se continúe estudiando a lo largo de casi un siglo, y a pesar de todo, parece que todavía se está lejos de llegar a comprender su naturaleza.

Cuando desde el punto de vista científico se trata de definir un concepto se busca que éste implique algunos elementos esenciales y estables que constituyan su núcleo teórico. Es el caso de la conceptualización y medición de la inteligencia, con sus múltiples puntos de vista, parece que dicho núcleo teórico "no puede ser delineado de una manera clara y aceptable. Se da tal falta de acuerdo sobre su significado, así como de sus aspectos operativos, teóricos y metateóricos, que el concepto de inteligencia no constituye un único constructo teórico" (Baltes en Stenberg, 1988, p. 40) de ahí que lo que podamos entender por inteligencia se convierta en un conjunto confuso de conceptos, ideas y problemas que en un momento se conectan pero en otro parecen irreconciliables.

Pero tratemos de entender qué es lo que ha sucedido alrededor de este fenómeno psicológico y tratemos de acercarnos lo más posible a los diferentes

enfoques y conceptos que sobre el tema se han desarrollado. Si bien la literatura que existe al respecto es basta, en este capítulo se tratará de recoger los principales estudios que se han realizado partiendo de aquellos trabajos que abordan desde su origen y/o aparición, hasta posturas que tratan de concebirla como un proceso básico, sin pretender llegar a la unificación de dichas posturas se tratará de abordar los diferentes conceptos y teorías, desde algunas perspectivas o corrientes generales.

## **2.1 Origen de la inteligencia**

Antes de definir lo que es inteligencia y sus diversas acepciones se considera conveniente iniciar este apartado con el trabajo realizado por William H Calvin (1994) en el que se aborda más que su conceptualización, su origen a partir del proceso evolutivo de ser humano.

Para la mayoría de los observadores la esencia de la inteligencia reside en la lucidez, la versatilidad para resolver problemas nuevos, también se considera que la esencia de la inteligencia es la previsión. "Piaget insistía en que la inteligencia era el complicado rodeo que damos cuando no sabemos que camino tomar; Barlow por su parte indica que la inteligencia es todo aquello que nos permite idear una conjetura que descubra un nuevo orden subyacente, esta definición cubre un aspecto muy amplio desde la solución de un problema, hasta dar con una analogía apropiada, una armonía agradable o una réplica ingeniosa" (Calvin. 1994, p79). La inteligencia como se ve implica aspectos tan elevados de nuestra vida mental que en determinado momento llega a reducirse a procesos tan elementales como por ejemplo; reconocer una calle o desbrocharse los botones de

la camisa y es precisamente en estos aspectos de la vida cotidiana donde se puede encontrar el punto de partida para explicar cuál su origen; o a partir de cuándo evolucionaron nuestras capacidades

La respuesta ha dichas interrogantes se podría encontrar en la teoría de la evolución y la neurofisiología, desde el punto de vista de esta última, la pregunta se centraría en las características que tiene nuestro cerebro, a esa parte del cerebro diría el autor, implicada en la producción de asociaciones nuevas en la corteza cerebral, de dos milímetros de espesor. Podría ser esta una de las explicaciones, del por qué somos más inteligentes que los animales, por lo que entonces una de las respuestas que se le puede dar a lo planteado es que la inteligencia surgió en virtud del refinamiento de alguna especialización del cerebro, tal como la que se da en el caso del lenguaje. La especialización sería entonces el salto fundamental en la evolución humana desde el tronco primate, hasta el surgimiento del hombre como tal, "esta especialización supone la existencia de una configuración física común para el lenguaje, la planificación de los movimientos de las manos, la música y la danza, entonces encierra mayor potencia explicativa" (p 79).

De lo anterior podrían desprenderse algunas características de la inteligencia: **la rapidez y la capacidad para combinar diversas ideas a la vez; la versatilidad**, generada por la posibilidad de tener una larga vida y dentro de ella un amplio período juvenil, lo que permite descubrir nuevos comportamientos, aunado a ello la vida social de los grupos humanos; **el lenguaje** por su parte "es lo más característico de la inteligencia humana; sin la sintaxis, seríamos poco más avispados que los chimpancés (...) La explicación del por qué los seres humanos

son tan inteligentes se puede encontrar en el lenguaje, sobre todo a partir precisamente de la sintaxis, y su capacidad para juntar sonidos sin significación, para hacer de ellos palabras significativas".- Otro aspecto y muy afín a la sintaxis es la **capacidad de hacer planes**. El ser humano a diferencia de los animales tiene esta capacidad de planeación – ( . . ) "la capacidad humana de hacer planes quizá se derive de nuestro talento para construir, mediante nexos sintácticos estructuras conceptuales más largas que las frases sencillas. Nuestra capacidad de hacer planes se desarrolla gradualmente a partir de los cuentos de la infancia y constituye un importante fundamento para las decisiones éticas; al imaginarnos el curso de acción, imaginamos a la vez los efectos y decidimos entonces actuar o no. De este modo la sintaxis eleva la inteligencia a un nuevo nivel" (Calvin,1994. p.80-81).

Pero ¿qué es lo que en realidad generó el desarrollo de la inteligencia? ¿depende de algún dispositivo nuclear de nuestro cerebro, de una configuración o plantilla que sirva para el lenguaje, para contar historias, para hacer planes, para los juegos y para la ética? Al respecto de estos cuestionamientos Calvin sugiere que las respuestas podrían encontrarse en "la planificación cerebral de los movimientos balísticos ....(acciones rapidísimas de los miembros que una vez iniciada, no se puede modificar. Por ejemplo, la de golpear un clavo con el martillo) . . pudo haber promovido el lenguaje, la habilidad musical y la inteligencia" (p. 82).

Siguiendo sobre esta línea de análisis, Secadas (1995) plantea que la conducta inteligente puede interpretarse a partir de la forma en que el hombre traduce la realidad, con determinados propósitos, así por ejemplo, en la actividad

cotidiana que realiza el pastor al cortar una rama para hacerse de un bastón se pueden distinguir cuatro formas de actividad transformadora:

**"Neótica:** generaliza la percepción en forma de concepto: cayado, largo, resistente constituye una respuesta común, abstracta o concreta pero no forzosamente verbal. Esta respuesta común es justamente la categorización del percepto, que transforma los datos perceptivos en un esquema conceptual abstracto

**Práctica o utilitaria :** el cayo o bastón sirve y se emplea para golpear, apalancar, sostener el cuerpo... Está pensando en función de usos

**Lógica Simbólica:** sirve como término de operaciones mentales: se toma como unidad de medida (una vara), simboliza la potencia, el mando su origen acompaña y sustenta el pensamiento, la mano de un obispo es símbolo pastoral.

**Ingeniosa o creativa:** interviene como artificio de creación de nuevas estructuras o de efectos originales: entre dos piedras hace un puente entre dos rueda un eje, etc." (Secadas, 1995 p. 25).

En síntesis el autor presenta un esbozo muy general sobre el desarrollo de la inteligencia, teniendo como punto de partida la acción que el hombre ejerce sobre el medio, al igual que Calvin, aunque no explícitamente, reconoce la existencia de procesos neurofisiológicos para comprender lo que es la inteligencia y su origen. Estas teorías nos remiten a una segunda pregunta y por tanto al siguiente apartado de nuestro capítulo.

## 2.2 ¿Cómo se define la Inteligencia?

### *Antecedentes*

Durante muchos años se ha tenido la idea de que las personas presentan variaciones en cuanto a lo que se llama inteligencia, en la antigüedad hace más de 2000 años, Platón estudió diversas variaciones sobre el concepto haciendo "hincapié en la existencia e importancia de los poderes mentales: las capacidades

que se han llamado en forma indistinta la racionalidad, inteligencia o el despliegue de la mente ( . ) Las expresiones "Conócete a ti mismo" de Sócrates; 'Todos los hombres por naturaleza desean saber', de Aristóteles, y "Pienso luego existo" de Descartes, constituyen epígrafes que enmarcan toda una civilización. La ascendencia de los factores intelectuales rara vez fue puesta en duda incluso en ese milenio oscuro entre los tiempos clásicos y el Renacimiento" (Gadner, 1995 p. 37).

Para los filósofos medievales, la inteligencia era la propia acción del intelecto; el proceso de conocimiento e interiorización de la realidad. Para Terman era la capacidad de razonamiento abstracto; para Thorndike, la capacidad de respuesta acertada que resulta posible por el ejercicio de multitud de funciones distintas que constituyen esa capacidad. De forma radical y graciosa, Boring negó que fuera posible plantear definición alguna y consideró que la inteligencia era simplemente "lo que miden los tests de inteligencia" (Anastasi, 1977), definición fundamentada obviamente en el criterio de validez de constructo de dichas pruebas.

Teorías más recientes acerca de su naturaleza básica "implican uno o más de los tres temas siguientes: la capacidad de aprender; el conocimiento total que una persona ha adquirido; y la capacidad de adaptarse a nuevas situaciones y al entorno en general" ( Woolfolk, 1996. p. 109).

Estas teorías las podemos encontrar expuestas en lo general en dos eventos de relevante importancia llevados a cabo el siglo pasado y son; el panel de expertos organizado por los editores de la revista "Inteligencia y su medición"

en 1921 y el organizado por Stenberg y Detterman en 1986, que en su momento trataron de definir la naturaleza de la inteligencia, sin llegar a un acuerdo en cuanto a su definición, es decir: a considerarla como una sola capacidad o muchas capacidades separadas, dichos eventos son el punto de partida para el análisis y discusión sobre el tema.

Lo anterior nos llevaría a plantear si la inteligencia representa una capacidad o muchas, las que plantean una capacidad se pueden considerar como las teorías clásicas, las que plantean muchas estarían representadas por las teorías modernas. Iniciamos entonces con las definiciones que podríamos considerar como clásicas para posteriormente pasar a aquellos planteamientos que encierran conceptualizaciones más amplias.

Para iniciar, diríamos que inteligencia viene de **Intus legere, ver dentro, como intuición de intus ire, ver el interior, penetrar.** Partamos de las concepciones planteadas en el Simposio titulado "La inteligencia y su medida", organizado por los directivos del *Journal of educational psychology*, en 1921, en donde se obtuvieron más de 14 definiciones sobre la inteligencia, entre las que se encuentran:

1. " La capacidad de dar respuestas que son ciertas u objetivas (E.L. Thorndike).
2. La capacidad para desarrollar el pensamiento abstracto (L.M. Terman)
3. La capacidad de aprender adaptarse al medio (S.S. Colvin).
4. La capacidad de adaptarse a situaciones reales relativamente nuevas (R. Pintner)
5. La capacidad de adquirir conocimientos, y los conocimientos que se poseen (V.A.C. Henmon)
6. Un mecanismo biológico por el que los efectos de una complejidad de estímulos son presentados al unísono, dando lugar a algún tipo de efecto unificado en la conducta (J.Peterson)

7. La capacidad para inhibir un ajuste instintivo, redefinir el ajuste instintivo inhibido a la luz de un proceso imaginario de tanteo, acierto-error; y aplicar el ajuste instintivo modificado en la conducta pública para la afirmación del individuo como animal social (L.L.Thurstone)
8. La capacidad de adquirir capacidades (H Woodrow)
9. La capacidad para aprender o sacar provecho de la experiencia (W.F. Dearbon)" (Stemberg y Detterman, 1986, p 8).

Vernon (1980) plantea que la inteligencia ha sido conceptualizada en diversos sentidos, en primer lugar sitúa el concepto que lleva implícito el significado de capacidad innata, algo que es heredado por nuestros antepasados a través de los genes y que determina el desarrollo mental del sujeto. En segundo lugar encontramos aquellos planteamientos que indican que el concepto inteligencia está referido para describir a un sujeto que es listo, rápido, tenaz, etcétera, es decir, que es mentalmente eficaz; un tercer significado es el que se relaciona con la edad mental del sujeto o puntuación en CI. (Coeficiente Intelectual) en cualquier test de inteligencia.

Secadas (1995) citando a varios autores define a la inteligencia de diversas maneras:

"Para G.E, Müller. Inteligencia es el esfuerzo dirigido a avanzar en el campo de las ideas; Köler habla de intuición, Yerkes se inclina por el pensamiento abstracto de medio y fin; Spencer prefiere la adaptación consciente; para Stern, es la adaptación a situaciones nuevas; Terman la define como pensar abstracto y ajuste a la verdad, Thurstone la entiende como resolución de alternativas y opciones incompletas, y como facilidad de aprender, para Wechsler sería capacidad global del individuo de obrar conforme a un propósito, pensar racionalmente y actuar con éxito en el entorno y para Binet la inteligencia exige cuatro capacidades fundamentales: comprensión, invención, dirección y crítica" (p. 513).

Hebb, distingue una inteligencia A, la cual indica el potencial en donde se elaboran formas inteligentes de actuar, implica básicamente ingénita y biológica. La inteligencia B o funcional, consiste en la eficacia de los recursos disponibles en

un momento dado y vendría determinado por el conjunto de estrategia aprendidas en interacción con el medio ambiente. Ambas formas las encontramos también en el concepto de Cattell quien distingue una inteligencia fluida que opera en las situaciones nuevas (A) y otra cristalizada en hábitos aprendidos de pensar (B).

Concepciones más complejas las encontraríamos con los siguientes autores Rouberoux y Carlier (1997) consideran que "inteligencia es una aptitud mental muy general que implica entre otras cosas, la capacidad de razonar, de prever, resolver problemas, pensar en abstracto, captar ideas complejas, aprender rápidamente y aprovechar la experiencia. No es solamente una capacidad personal, una facultad estrictamente académica o una aptitud para los test. Refleja una capacidad más amplia, profunda de comprender nuestro entorno de 'asir al vuelo', de dar sentido a las cosas y de imaginar soluciones prácticas" (en Secadas, 1995. p.265).

Secadas (1995) por su parte plantea que obramos inteligentemente cuando somos capaces de resolver las situaciones que se nos plantean o nos planteamos nosotros mismos. El ser inteligente implica disponer de recursos para resolver problemas prácticos, aprender fácilmente, estar capacitado para entender y pensar, para el autor "el pensamiento representa la suprema integración de la conducta donde confluye a manera de esquemas condicionantes; la percepción, el aprendizaje, la motivación y los dinamismos sociales y axiológicos de la experiencia", ( p. 511) es decir, en todas nuestras actividades hacemos uso de la inteligencia, inclusive en aquellas que podríamos considerar vanas, como el aburrimiento y la ociosidad. Para el autor referido representa "la capacidad de comprender y manejar la realidad, resolver situaciones complejas, anticiparse a los

efectos prefigurados en la situación, ampliar los dominios del saber y mantener una dirección concertada y eficaz de comportamiento, merced a una selección depurada de experiencias, hábitos, recursos, y a un flexible replanteamiento de coyunturas, disposiciones y actitudes" (p.513).

Por otro lado Pastor y Sastre (1994, p. 193), adoptan una concepción basada en el desarrollo cognitivo, y consideran a la inteligencia como "la capacidad de establecer relaciones que presupone la posibilidad de estructuración y/o categorización del saber, y el desarrollo cognitivo como la progresiva realización y actualización (o reestructuración) de dicha capacidad". Lo anterior en su conjunto supone de acuerdo con los autores una adaptación e interacción continua (en feed-back) al medio físico y social

John Carroll (1988, p. 68), plantea que la inteligencia debe ser considerada en general como un concepto en la mente de una sociedad, que tiende a caracterizar a los sujetos de acuerdo a la aparente capacidad que éstos poseen, para enfrentarse a diversas situaciones y tareas, en tres campos problemáticos específicos: a) académicos y técnicos, b) prácticos y c) sociales. "Los problemas académicos y técnicos son aquellos que surgen en el trabajo escolar y, más tarde, en muchos campos de la ciencia, de las profesiones y de las especialidades ocupacionales. Los problemas prácticos son los que se plantean al afrontar los asuntos cotidianos, al ganarse la vida y al planificar el curso de nuestras propias acciones, incluso al planificar la propia vida. Los problemas sociales son los que surgen en la interacción con los otros-individuos, grupos o incluso culturas extrañas". Por tanto considera que no podemos hablar de una sola concepción sino de diferentes clases de inteligencia, académica, práctica y social, y que

dependiendo de los tipos de problemas a los que nos enfrentemos serán aplicadas con mayor o menor éxito, dependiendo también de la rapidez de la ejecución se nos considerará más capaces e inteligentes. Lo cierto es que todas estas concepciones encierran dentro de sí ciertos mitos y verdades veamos que nos dice Sternberg (en González, 2000) al respecto.

### **2.2.1 Mitos y verdades acerca de la inteligencia**

Como se ha venido abordando a lo largo de este capítulo, "el tema de la inteligencia ha tenido un fuerte impacto en los especialistas quienes han producido nuevos y valiosos conocimientos sobre la inteligencia humana pero también se han creado verdades a medias y mitos que poco han sido cuestionados. En un intento por aclarar algunas de estas concepciones, especialmente a partir de la publicación del libro *The bell curve* de Herrnstein y Murray (1994), Sternberg (1996) realizó un interesante trabajo en el cual identificó los diferentes mitos, contra mitos y verdades que han surgido en torno a la inteligencia que yacen tanto en la citada obra como en muchas de las creencias populares de este complejo concepto" (Gonzalez,2000, p. 15) veamos la siguiente tabla:

**Tabla 1. Mitos y verdades sobre el concepto de inteligencia.**

Mito	Contra mito mítico	Verdad
1. La inteligencia es una cosa, "g" (C.I.)	La inteligencia es muchas cosas que difícilmente pueden contarse.	La inteligencia es multidimensional pero científicamente manejable.
2. El orden social es un resultado natural en gran parte del orden del C.I.	Los tests crean totalmente un orden social.	El orden social es parcialmente pero no exclusivamente creado por los tests.
3. La inteligencia no puede ser enseñada en ningún grado significativo.	Podemos lograr increíbles hazañas en la enseñanza a los individuos para ser más inteligentes	Podemos enseñar inteligencia al menos en cierto grado pero no podemos efectuar cambios radicales en este aspecto.
4. Los tests de C.I. virtualmente miden todo lo que es importante para el éxito en la escuela y el trabajo.	Los tests de C.I. no miden virtualmente nada importante para el éxito escolar o laboral	Los tests de C.I. miden habilidades que tienen una moderada importancia en el éxito escolar y laboral.
5. Estamos utilizando muy poco los tests perdiendo información valiosa	Estamos utilizando excesivamente los tests los cuales deberían ser abolidos.	Los tests, cuando son adecuadamente interpretados, pueden tener una función útil pero limitada. Pero sucede que frecuentemente no son adecuadamente interpretados.
6. Nosotros, como sociedad, estamos siendo cada vez más tontos debido a los efectos de la multiplicación de tontos genéticos.	No tenemos ninguna razón para temer una disminución en las habilidades intelectuales entre las sucesivas generaciones.	Tenemos algunas razones para temer una pérdida de habilidades intelectuales en las futuras generaciones pero el problema no se debe a una sobreproducción de tontos
7. La inteligencia es esencialmente toda heredada con excepción a una cierta varianza trivial e inexplicable.	La inteligencia es totalmente ambiental con excepción a una varianza trivial e inexplicable.	La inteligencia comprende componentes heredables y ambientales en interacción
8. Las diferencias raciales en C.I. conducen claramente a las diferencias en los resultados	Las diferencias raciales en C.I. no tienen nada ver con las diferencias en el ambiente.	Aún no comprendemos realmente las relaciones entre raza, C.I. y ambiente.
9. Deberíamos expulsar a la gente tonta.	No hay cosas tales como una persona tonta. Todos son listos.	Necesitamos repensar lo que entendemos por "tonto" y "listo".

Tomado de Sternberg (1996, p. 12, traducción de González, 2000, p.15)

En la tabla se observa que el concepto que sobre la inteligencia se tiene, puede ser tan amplio y complejo; simple y sencillo, según la idea que sobre ella se tenga tanto por parte de los expertos como de los no expertos. Éste tenderá a variar por su contenido, según subrayen la capacidad de adaptación, la facilidad para aprender, la actitud para el pensamiento abstracto y la resolución de problemas. Se puede considerar que el ser inteligente, es disponer de recursos para resolver tanto problemas prácticos, como el estar capacitado para entender y pensar. Por lo tanto hablar de inteligencia puede ser tan sencillo o complejo según lo queramos entender, lo que sí es claro es que las definiciones dadas (algunas vagas, muy generales o más complejas) son solamente un principio para entender lo que es la inteligencia, ya que por sí mismas "no ofrecen ninguna garantía de que las nociones subjetivas de las personas sean representaciones verídicas de sus capacidades intelectuales. En consecuencia, estas definiciones se entenderán

mejor, probablemente, como explicatorias de la 'situación actual' y necesitadas de ser completadas por más teorías científicas (formales) de la naturaleza de la inteligencia" (Sternberg, 1986. p 10). Lo que nos remite a nuestro siguiente planteamiento.

### 2.3 Teorías que explican la inteligencia

Debido a que la inteligencia no puede ser restringida a una aptitud más, los estudios que al respecto se han desarrollado, han llevado desde los trabajos pioneros de F J. Galton, A Binet, Ch. Spearman y L.L Thurstone, a la elaboración de una serie de modelos y teoría que han girado en torno a una dicotomía: la de considerar a la inteligencia como una estructura unitaria o bien como una estructura múltiple. Durante todo un siglo este debate ha estado abierto y sólo recientemente hay un acercamiento entre ambas posturas. "Estas formas de modelizar la inteligencia han sido aceptadas por distintas aproximaciones psicológicas, así la concepción unitaria es más propia de la tradición diferencialista de la psicología y la segunda más habitual en el contexto profesional de la psicología" (Pueyo, 1999, p 2)

El acercamiento entre estos dos planteamientos ha generado en los últimos años que los mismos factorialistas la consideren como algo modificable a partir del genotipo inicial, y con ello a diferentes clasificaciones de la inteligencia en; **Inteligencia A** fundamentalmente referida a la estructura y funcionamiento cerebrales, una **Inteligencia B** o inteligencia funcional, y una **Inteligencia C**, la inteligencia ecológica, por la que los sujetos resolvemos los problemas que nos plantea nuestra cultura y nuestra vida cotidiana. (Mora, 1993, p 286)

La integración de estos tres tipos de inteligencia se pueden ver afectados por la historia de vida de los sujetos. es decir desde el punto de vista de la psicología diferencial, es útil el estudio de la inteligencia en cuanto a su repercusión para el surgimiento de las diferencias entre los sujetos y grupos humanos; diferencias en función del sexo, interrelaciones y correlaciones entre los diversos tipos de inteligencia antes mencionados, las relaciones de dotación genética con la inteligencia, las variaciones instructivas y su repercusión en la inteligencia, las diferencias entre los diversos grupos humanos, las diferencias de personalidad, los estilos cognitivos, etc.

Todos estos aspectos nos llevan a considerar en términos generales el planteamiento de que todos los estudios que sobre inteligencia se han realizado, suelen versar básicamente sobre cinco aspectos interdependientes: "estudios en los que se busca establecer ¿qué es o en qué se hace consistir?, a través de definiciones y compilaciones; el aspecto procesual: ¿cómo funciona?, o aquellos que plantean ¿de qué forma se desarrolla?. Su aspecto evolutivo; Diferencial ¿cómo se diferencian unos individuos de otros?. Y finalmente su aspecto aplicado en el que se trata de establecer ¿cómo puede el individuo beneficiarse de tales diferencias?" (Secadas, 1995, p 512).

Estas interrogantes nos permiten ubicar el proceso que ha seguido el estudio de la inteligencia y por tanto los diversos planteamientos teóricos, así como los modelos que han surgido al respecto dentro del ámbito de la psicología, las teorías que explican la inteligencia han sido desarrolladas principalmente por la tradición psicométrica por un lado, y por las del procesamiento de información, los

cuales no son excluyentes una de la otra y si complementarias. De acuerdo con Sternberg (1990) estas teorías se encuentran enmarcadas en lo que se ha definido como teorías explícitas de la inteligencia que implican aquellas descripciones formales elaboradas por los psicólogos, que conforman la base de la mayor parte de las investigaciones empíricas, de aquí su importancia ya que son éstas las que dan fundamento a la elaboración de instrumentos para la medición de la inteligencia

Para clarificar en que consiste cada una de ellas retomaremos las siguientes definiciones: "*Las teorías psicométricas o diferenciales* se centran principalmente en la cuestión de qué constelaciones estructurales de procesos, contenidos, representaciones dan lugar a modelos estables de diferencias individuales. *Las teorías del procesamiento de información* (también conocidas como cognitivas) intentan identificar realmente los procesos, contenidos y representaciones individuales". (Sternberg, 1990, p.26-27)

Por su parte Yela (en Mayor y Pinillos, 1991) es uno de los investigadores que han tratado de establecer las relaciones que se dan entre dos disciplinas de la psicología; la experimental y la correlacional, indicando que ambas se encuentran presentes en el estudio de la inteligencia, "estas dos grandes corrientes metodológicas se han aplicado al examen de tres grandes cuestiones: la general, la diferencial y la genética. La psicología general de la inteligencia trata de averiguar las regulaciones, normas y leyes que rigen su estructura y dinamismo. La diferencial de descubrir las aptitudes y los componentes que la constituyen y en los que los individuos y los grupos difieren. La genética de indagar el origen y

desarrollo de los componentes, como procesos, las estructuras y las aptitudes’  
(p.286)

Como se puede ver son diversos los modelos y las teorías que tratan de entender la estructura y organización de la inteligencia encontrando puntos de coincidencia entre los autores citados, en cuanto a la existencia de dos tendencias claramente definidas, las cuales podríamos considerar como aquellas visiones clásicas, encontrando a su vez dos tendencias, **la multifactorial**; que considera “que hay tantas inteligencias como situaciones o exigencias ambientales puede encontrar el hombre en su vida cotidiana, y estas inteligencias son independientes entre sí” –comporta un conjunto de aptitudes específicas, sin una jerarquía, actuando en función de las situaciones a que se enfrenta el sujeto. La segunda que considera que la estructura de la inteligencia esta organizada en forma de - “una **estructura jerárquica o unitaria** en la cual las aptitudes están interrelacionadas y son dependientes unas o varias capacidades más amplias, a las que están subordinadas”, éstas han seguido la tradición iniciada por Spearman y desarrollada por autores como R B Cattell y otros. (Corriente que será retomada con más detalle en la parte central de este capítulo, ya que dentro de ella se encuentra el sustento teórico del test de factor “g” de Cattell, objeto central de este trabajo). Estas teorías constituyen los principales modelos empíricos de la inteligencia y ambas comparten un enfoque metodológico común, el análisis factorial, constituyendo así la llamada tradición psicométrica del estudio de la inteligencia, asociadas a la justificación de su medición y especialmente al Cociente Intelectual (Pueyo, 1999, p. 2-3)

Tratemos entonces de explicar cada una de ellas considerando sus antecedentes y principales planteamientos

### ***2.3.1 Teorías psicométricas o diferenciales de la inteligencia humana***

Uno de los principales antecedentes que podemos encontrar para el desarrollo no sólo del estudio de la inteligencia sino de otras áreas de la ciencia es el libro el origen de las especies de Charles Darwin, en el cual "se proponía que la evolución de las especies y el desarrollo de los seres humanos podría remontarse a un proceso evolucionista de la selección natural (...) el libro sugería que, en cierto sentido, las capacidades de los seres humanos no presentaban solución de continuidad respecto de las capacidades de los animales inferiores" (Stenberg, 1997, p 55) lo que generó interrogantes respecto a la relación existente entre las capacidades de los animales inferiores y el hombre. Galton primo de Darwin fue uno de los primeros en investigar las implicaciones del libro en el campo de la inteligencia humana, de sus estudios desprendió dos cualidades que a su parecer permitían distinguir a la gente más inteligente de la que lo es menos; la primera es la energía o capacidad para el trabajo; la segunda la sensibilidad. Estas ideas son posiblemente el punto de partida para la aproximación de los psicólogos al estudio de la inteligencia y que éstos fueran analizando las diferencias entre diversos grupos y entre los diferentes sujetos humanos.

Partiendo de lo anterior podríamos decir que las teorías psicométricas encuentran su fundamento en el estudio de las diferencias individuales que existen entre las personas, pero especialmente de un grupo de habilidades subyacentes esenciales, como la verbal, de razonamiento, memoria, etc., las cuales son

identificadas mediante una técnica matemática llamada análisis factorial. Esta técnica comienza con una matriz de intercorrelaciones (o covariaciones) para una serie de test e identifica el origen 'latente' de la variación que subyace en la puntuación de las pruebas, (...) de este modo se propone que las diferencias individuales en la realización de las pruebas de inteligencia se pueden descomponer en diferencias individuales en estos factores, y se postula que cada uno de ellos representa una habilidad humana distinta". (Stenberg, 1990. p 26)

Dado que el manejo de factores es un elemento común en estas teorías para entender la inteligencia, es necesario diferenciarlas, para determinar con mayor precisión sus características. Son dos las principales diferencias que vamos a localizar. La primera se basa en el número de factores propuesto por la teoría; segundo, por la disposición geométrica de cada factor respecto a otro. Veamos cada una de éstos.

**a) Variaciones en el número de factores** Estos pueden ir de 1 a 150 considerando las teorías principales. Iniciando con Spearman (1927) plantea que son dos clases de factores que explican la inteligencia: un primer factor de tipo general que impregna la realización de todos los ejercicios intelectuales; y factores específicos para cada una de las habilidades que impregna un solo único ejercicio. Para el autor el primer planteamiento es el que adquiere mayor interés psicológico y es el factor general o "g", a partir de la cual generó dos propuestas respecto a su naturaleza: "Una sugería que las diferencias individuales dentro de "g" pueden ser entendidas en términos de diferencias en los niveles de energía mental que los individuos podrían aportar en la realización de una labor intelectual.

La otra propuesta era que las diferencias individuales en 'g' podrían ser consideradas como diferencias en las aptitudes de las personas para utilizar tres principios cualitativos de cognición: percepción de la experiencia educación de las relaciones, y educación de elementos correlacionados" ( Stenberg, 1990, p. 27).

En la analogía Abogado- Cliente= Doctor: ?; se puede observar la aplicación de estos tres principios, así, al codificar los términos dados en la analogía, estamos aplicando el primer principio, ya que al percibir y comprender los conceptos estamos apelando a la percepción En la analogía: **Abogado - Cliente =Doctor - ?**; se puede observar la aplicación de estos tres principios, así, al codificar los términos dados en la analogía, estamos aplicando el primer principio, ya que al percibir y comprender los conceptos estamos apelando a la percepción de la experiencia; al establecer la relación entre Abogado y Cliente estamos recurriendo al segundo principio, educación de relaciones; y finalmente al aplicar la regla deducida de Abogado a Cliente, la respuesta analógica para Doctor sería Paciente, por lo que recurrimos a la educación de elementos correlacionados. Lo anterior llevó a Spearman y a muchos otros investigadores a plantear que el pensamiento de tipo analógico es una de las mejores medidas disponibles para explicar la inteligencia de factor "g", y a muchos otros investigadores a plantear que el pensamiento de tipo analógico es una de las mejores medidas disponibles para explicar la inteligencia de factor "g" y a cuestionar las doctrinas monárquicas de las aptitudes cognoscitivas, la, cual dado que define la existencia de una sola facultad cognoscitiva, cosa que experimentalmente no se encuentra (Stenberg, 1990, p. 27).

La misma crítica la realiza el autor a las doctrinas anárquicas la cual plantea lo contrario; la existencia de una multiplicidad inmensurable de factores elementales e independientes en donde contrario a lo que se plantea, si se presentan en los trabajos experimentales correlaciones positivas en las aptitudes aunque no perfectas, se ordenan jerárquicamente. Sin embargo, también considera que ambas teorías tienen algo de verdad "No hay una sola facultad cognoscitiva, pero hay un solo factor común a todos los procesos gneseológicos Y si no hay únicamente innumerables factores elementales, si hay en cambio, tantos factores específicos como opciones distintas, factores que, junto con lo general, constituyen las innumerables aptitudes específicas" (Cerde, 1973, p 131).

## b) Variaciones en la estructura geométrica de los factores

De acuerdo con Stenberg (1990, p.28-29) son cuatro las estructuras más conocidas para determinar las diferentes disposiciones geométricas de los factores dispuestos para definir la inteligencia:

**1. Disposiciones jerárquicas.** En términos generales las estructuras jerárquicas son las que más se han utilizado en los trabajos sobre psicología diferencial, es a partir de este punto de vista que se ha llegado a determinar que las aptitudes presentan diferentes niveles de importancia, ya que algunas son más globales que otras y por tanto más importantes, un ejemplo de ello sería el modelo factorial de Spearman con un factor general y otros factores específicos de menor importancia.

**2. Las disposiciones no jerárquicas.** Consiste en no colocar en lugar especial alguno de los factores seleccionados, presentados en forma de lista a todos se les concede la misma importancia sin considerar el lugar que ocupen, la puntuación obtenida en un factor es igualmente válida para otro. Un ejemplo de este tipo de estructura sería las siete capacidades mentales primarias de Thurstone.

**3. Disposición cúbica.** Uno de sus principales representantes es Guilford (1982) él representó la estructura del intelecto como un gran cubo compuesto por 150 cubos pequeños. Cada dimensión del cubo corresponde a una de las tres categorías (operación, contenido y producto), y cada una de las 150 combinaciones posibles de las tres categorías forman uno de los cubos pequeños.

**4. La estructura "radex"** para la inteligencia; considerado el radex como un círculo, Guttman (1965) propone que los diferentes test que se encuentran dentro de una prueba de inteligencia pueden situarse en algún lugar dentro del círculo. Los test que están más centrales para la inteligencia, por lo que estas se consideran como las medidas más puras y las periféricas las menos puras.

Dentro de esta postura encontramos también el manejo de otros enfoques planteados por Castelló (en Genovard, Beltán y Rivas 1995), a saber; los enfoques monolíticos o unifactoriales, fundamentalmente representados por los modelos del

CI. y factor "g", el desarrollo de estos modelos han encontrado una estrecha relación entre la inteligencia y las actividades académicas, destacándose el uso de test unifactoriales ligados al CI. Para explicar con mayor claridad como ha sido establecida esta relación por varios autores, se citaran algunos elementos:

### **2.3.1.1 Modelos monolíticos o unifactoriales de la inteligencia**

- a) Los modelos monolíticos de la inteligencia utilizaron como principal parámetro de validación la competencia académica de los sujetos, fuera directamente o a través de consecuencias de la misma (nivel de estudios o acceso a profesiones que precisaban de estudios universitarios, por ejemplo)
- b) Los instrumentos de medida derivados de dichos modelos tenían como función principal la discriminación entre estudiantes eficaces y estudiantes lentos o ineficaces. En el caso de Binet, esta función es completamente explícita, al igual que los instrumentos de stanford-Binet, Terman y Weschler, donde la validación concurrente complementaba la validación a partir de la eficacia académica de los sujetos.
- c) Las únicas predicciones, realizadas con los test unifactoriales de este tipo, que han mantenido cierta estabilidad y han soportado los análisis metodológicos, han sido las referidas al rendimiento académico.
- d) Los elementos de información culturalmente determinados o ligados a la escolarización forman, habitualmente, parte de estos tres. Por consiguiente evaluar un CI, significa tanto evaluar capacidades de procesamiento como informaciones adquirida (Castelló en Genovard, Beltán y Rivas, 1995, p.204)

Estos cuatro aspectos apoyan de manera directa aquellos planteamientos en los que la inteligencia es evaluada a través de los test del CI, tal como lo plantea el modelo se corresponde con la inteligencia académica

En cuanto a los modelos del factor g, se incluyen también dentro de las teorías monolíticas, ya que considera a la inteligencia como una sola variable y su validación se da en los ámbitos académicos, sin embargo, se pueden identificar con respecto al planteamiento anterior, "la metodología utilizada en los modelos del factor g, es el principal esqueleto teórico de los mismos (...) el método

matemático de análisis factorial (especialmente el centroide y no rotado) aplicado a los test intelectuales existentes a mediados de los años veinte, genera la identificación de los elementos comunes de los mismos, aglutinados en el primer factor centrado. constituyen la base del concepto de inteligencia general (factor g)" (Castelló en Genovard, Beltrán y Rivas, 1995, p. 206).

La eliminación de los factores culturales o ligados a la información es otro elemento que caracteriza a los modelos del factor g, encontrándose como una constante entre los investigadores, con la idea de mantener los aspectos más puramente procesuales, aportando con ello que las aptitudes vinculadas a los procesos lógicos, a la inducción y a la generalización constituyan el núcleo de los instrumentos de factor g. Los trabajos realizados por Anastasi, Eysenk y Raven, son ejemplos de depuración de las influencias de tipo verbal y sociocultural, así como la mínima intervención de la memoria, que buscan convertirse en medidas libres de cultura, donde el razonamiento lógico y los procesos inductivos-deductivos son los factores que más sólidamente aparecen representados

Los modelos anteriormente expuestos se podría considerar que son equivalentes tanto por su fundamento, como por su omisión de los aspectos más culturales

### **2.3.1.2 Modelos factoriales de la inteligencia**

Seguido a los modelos monolíticos desde un punto de vista histórico, encontramos los modelos factoriales de la inteligencia, que fundamentados en la metodología del análisis factorial iniciada por Spearman, cambia del método

centroide, al sistema de componentes principales se introduce la rotación de factores, se ajustan los valores matemáticos a características psicológicas o psicométricas. Respetando la base conceptual anterior los autores factoriales consideran a la inteligencia "como un conjunto de funciones mentales o aptitudes conductuales, las cuales, vistas globalmente definen el perfil intelectual del individuo(.....) Los valores de los componentes de dichos perfiles pueden diferir de un individuo a otro, generando configuraciones intelectuales diferenciadas" (Gardner en Castelló, 1995, p. 207). Este enfoque permite considerar la competencia intelectual de un sujeto desde un punto de vista más cualitativo, dejando atrás las interpretaciones meramente cuantitativas, asumiendo que existen formas diferenciadas de funcionamiento cognitivo, según la configuración intelectual. Ejemplos de ello lo podemos encontrar en "las configuraciones basadas en el pensamiento lógico o el creativo, o las diferencias tópicas entre las formas de funcionamiento más adecuadas para las letras o las ciencias (quizá más correctamente definidas como configuraciones basadas en el procesamiento verbal y holístico, en oposición a pensamiento cuantitativo y analítico)" (p.207).

Veamos algunos ejemplos de estas teorías:

Thurstone (1938), es uno de los primeros investigadores que desarrollan una concepción factorial de la inteligencia, quien postula que ésta es "un conjunto de habilidades mentales primarias sin ninguna preeminencia entre ellas" (Gadner, 1995, p. 41), define siete habilidades mentales esenciales. A continuación se describen:

a.- Comprensión verbal. Esta habilidad es típicamente medida mediante pruebas de vocabulario (incluyendo sinónimos y antónimos) y mediante pruebas de habilidad de lectura comprensiva.

- b.- Fluidez verbal. Se mide normalmente mediante test que requieren una producción rápida de palabras.
- c.- Habilidad numérica. Es típicamente medida mediante problemas aritméticos en los que hay cierto énfasis tanto en el cálculo como en el razonamiento, pero un énfasis relativamente pequeño en el alcance del conocimiento anterior.
- d.- Visualización espacial. Es típicamente medida mediante test que requieren la manipulación de símbolos o diseños geométricos.
- e.- Memoria. Es medida mediante una prueba de recuerdo-memoria de palabras o frases o mediante el recuerdo por asociación de parejas de nombres con fotos o imágenes de personas, con tiempos fijos.
- f - Razonamiento. Esta habilidad se mide típicamente mediante test como analogías y terminación de series.
- g.- Rapidez de percepción. Implica pruebas en las que se requiere el reconocimiento rápido de símbolos (Stenberg. 1990, p.27-28).

En estudios más recientes sobre los factores que definen a la inteligencia, encontramos a Guilford (en Sternberg, 1990), quien propone que la inteligencia abarca 150 factores distintos, según el autor, cada ejercicio o trabajo mental conlleva tres ingredientes: una operación, un contenido y un producto. Existen cinco tipos de operaciones: conocimiento, memoria, producción divergente, producción convergente y evaluación. Hay cinco tipos de contenido: visual, auditivo, simbólico, semántico y comportamentalista. Y existen seis clases de productos: unidades, clases, relaciones, sistemas, transformaciones e implicaciones. Puesto que las categorías se definen independientemente, son multiplicativas, así que hay  $5 \times 5 \times 6 = 150$  factores distintos. El autor plantea que estos factores lógicamente independientes, psicológicamente son dependientes ya que se encuentran intercorrelacionados. Esta teoría se conoce también con el nombre de estructura del intelecto.

En términos generales las estructuras jerárquicas son las que más se utilizan en los trabajos sobre psicología diferencial, es a partir de este punto de

vista que se ha llegado a determinar que las aptitudes presentan diferentes niveles de importancia, ya que algunas son más globales que otras y por tanto más importantes, un ejemplo de ellos sería el modelo factorial de Spearman (1927) con un factor general y otros factores específicos de menor importancia.

En síntesis la diferencia básica en las teorías psicométricas o diferenciales de la inteligencia se encuentra en la disposición de los factores, ya sea en función de su número o por su ordenación geométrica.

### ***2.3.2 Teorías cognitivas o del procesamiento de la información sobre la inteligencia humana***

Parten del planteamiento de considerar la función que tienen los procesos cognitivos en la búsqueda de respuestas por parte del sujeto para solucionar un determinado problema, y a partir de esto poder entender la inteligencia humana. Dentro de la literatura se han encontrado principalmente dos puntos de vista para explicar como se llevan acabo estos procesos: una primera postura que busca entender la inteligencia en función de la velocidad absoluta con la que un sujeto elabora la información percibida/recibida. En segundo lugar aquellas investigaciones que han estudiado formas más complejas utilizadas por el sujeto para la resolución de problemas. “En general, el mayor énfasis en la rapidez de la asimilación a sido asociado a los investigadores que estudian las formas más simples de elaboración de la información, mientras que el mayor énfasis en la exactitud y estrategias en la elaboración de información ha sido asociado con los investigadores que estudian las formas más complejas de dicho proceso”

(Sternberg, 1990. p. 31) De acuerdo con Sternberg (1990), algunos estudios realizados sobre ambos aspectos son los siguientes:

1. **Rapidez neta:** Plantea que las diferencias individuales en los sujetos puede estar dada por la rapidez absoluta en la que el sujeto tarda en elaborar la información, su paradigma se basa en el tiempo de reacción simple, dentro de éste, se le pide al sujeto que "dé una única respuesta abierta lo más rápidamente posible tras la presentación de un estímulo" (p.32). Entre sus principales usuarios se encuentra Galton (1883), Cattell (1890), Lemon (1927), Lunneborg (1977) y Jensen (1980).
2. **Rapidez de selección.** Parte de que la inteligencia no se deriva de una simple velocidad de elaboración, sino de la rapidez con que el sujeto selecciona o toma decisiones respecto al estímulo que se le presenta; dentro de este paradigma al sujeto se le presentan uno, dos o más estímulos, ante los cuales tiene que emitir una respuesta o reacciones diferentes, el propósito es que elija lo más rápidamente la respuesta correcta tras la presentación del estímulo. Los estudios más interesantes a este respecto han sido los elaborados por Jensen (1979; 82), Lally y Nettelbeck (1977), quienes descubrieron que la correlación entre el tiempo de reacción de selección y el Cociente Intelectual tiende a aumentar con el número de selecciones de estímulo-respuesta presentes en la prueba (p.33).
3. **Rapidez de acceso léxico.** Hunt (en Sternberg, 1990) planteó que las diferencias individuales en la inteligencia verbal tienen que ver con la capacidad que el sujeto tiene de la rapidez con que éste accesa a la información léxica mediante una memoria a largo plazo. A diferencia de las dos anteriores, las pruebas de acceso léxico ofrecen un panorama más amplio en cuanto al entendimiento de la inteligencia medida.
4. **Rapidez en los procesos de razonamiento** Investigadores como Pellegrino y Gardner (1983) y Whiteley (1980), basados en los principios de cognición cualitativos de Spearman, hacen hincapié en el papel que juega el razonamiento en la elaboración de la información por parte de los sujetos cuando se enfrenta a ejercicios tales como analogías, silogismos, terminaciones de series, entre otros, y que son considerados como trabajos de alto nivel de razonamiento. A partir de este planteamiento dichos investigadores se proponen conocer la inteligencia en función de dichos procesos de realización y de las habilidades para ejecutarlos rápidamente. Los investigadores que persiguen el conocimiento de la inteligencia en función de los procesos ejecutivos, intentan descubrir los procesos mediante los cuales los individuos toman decisiones tales como: a) qué componentes funcionales utilizan para resolver diferentes tipos de problemas, b) cómo combinar dichos componentes en una estrategia global, c) cómo representar la información, d) como intercambiar rapidez por exactitud, etcétera. Sternberg (1981, en Sternberg, 1990) ha propuesto que estos procesos ejecutivos o

'metacomponentes' son muy generales a los cometidos que implican un funcionamiento inteligente y que son en gran parte responsables de que surja un factor general en las pruebas de habilidad mental.

### **2.3.2.1 Teorías Múltiples de la Inteligencia.**

A partir del continuo estudio de las teorías psicométricas clásicas, se han desarrollado otros planteamientos que buscan romper con la ortodoxia de los modelos psicométricos de la inteligencia en los que ésta es considerada como una sola y que de ella se desprenden las habilidades generales o específicas de los sujetos, y en los que la base fundamental de su trabajo es el análisis factorial. Las aportaciones de dos importantes psicólogos norteamericanos R.J. Sternberg; profesor de la Universidad de Yale y H. Gardner de la Universidad de Harvard, proponen la existencia de múltiples inteligencias, las cuales intentan recoger los avances realizados por las ciencias cognitivas. "Según Gardner, los desarrollos recientes de los estudios de la inteligencia artificial, de la psicología del desarrollo, la neurología y la nueva ciencia cognitiva plantean la necesidad de superar el esquema clásico de la inteligencia como un concepto unitario y sustituirlo por un concepto multiforme de la misma" (Pueyo, 1999, p 4).

Pasemos a describir brevemente cada una de las propuestas de desarrolladas por estos autores:

### ***Modelo triarquico de la inteligencia humana de R.J. Sternberg.***

El modelo desarrollado por el autor pretende explicar la inteligencia a partir de tres tipos de inteligencia; la inteligencia analítica o componencial, la inteligencia experiencial y la inteligencia contextual. Estos tres tipos de inteligencia se ponen en juego continuamente y nos permiten confrontarnos a las situaciones tanto cotidianas, como nuevas en contextos distintos y aún desconocidos.

Su modelo triárquico consiste en una *Subteoría componencial* en la que establece cómo la inteligencia; refleja el mundo interno de cada uno a través de tres tipos de procesos: a) Metacomponentes o procesos de planificación, control y evaluación; b) Componentes de desempeño, que utilizamos para llevar a cabo las instrucciones de los procesos ejecutivos y c) Componentes de adquisición del conocimiento, que usamos para aprender, comprender y recordar información. A través de este tipo de inteligencia, se pretende obtener un diagnóstico sobre la capacidad del sujeto para pensar en forma abstracta, procesar información y determinar qué se necesita hacer. Las tareas que se utilizan para medir este tipo de inteligencia con base en los componentes son: analogías, vocabulario y silogismos (Wolfoolk, 1996, p.112)

*Subteoría experiencial.* Trata acerca de la relación entre inteligencia y experiencia, a través de ella el sujeto muestra su capacidad para formular nuevas ideas y combinar hechos no relacionados, basada en la experiencia, evalúa la capacidad de una persona para manejar tareas nuevas de manera automática.

*Subteoría Contextual.* Trata de la relación existente entre la inteligencia y el mundo externo del individuo, a través de ella se obtiene la información pertinente para determinar la capacidad del sujeto para adaptarse a un entorno cambiante y modelar el mundo propio para optimizar las oportunidades, prepararse para la solución de problemas en situaciones específicas. De acuerdo con Sternberg, la aplicación de este tipo de inteligencia al ambiente supone "la puesta en servicio de tres funciones de conducta inteligente: adaptación a los medios existentes, configuración de estos para transformarlos en nuevos entornos y selección de ambientes novedosos" (Wolfoolk, 1996, p. 113).

### ***El modelo de las inteligencias múltiples de Gardner***

Según Gardner existen siete inteligencias distintas que constituyen las formas como los individuos adquieren, retienen y manipulan la información del medio y demuestran sus pensamientos a los demás. Los estudios realizados por el autor permitieron diferenciar los tipos de inteligencia según las habilidades o destrezas cognitivas identificadas en diferentes poblaciones de sujetos particulares; talentosos, con lesiones cerebrales, observaciones evolutivas y transculturales. De los resultados obtenidos propuso su teoría de las siete inteligencias, a continuación se describen:

1. **Inteligencia lingüística.** Se utiliza en la lectura de libros, en la escritura de textos y en la comprensión de palabras y uso del lenguaje cotidiano. Esta inteligencia se observa en los poetas y escritores, pero también en oradores y locutores de los medios de comunicación.
2. **Inteligencia lógico-matemática;** utilizada en la resolución de problemas matemáticos, en el contraste de un balanceo o cuenta bancaria y en múltiples

tareas que requieran el uso de la lógica inferencial o proposicional. Es la propia de los científicos.

3. **Inteligencia musical:** se utiliza al cantar una canción componer una sonata, tocar un instrumento musical o al apreciar la belleza o estructura de una composición musical, naturalmente se observa en compositores y músicos en general
4. **Inteligencia espacial:** se utiliza en la realización de desplazamientos por una ciudad o edificio, en comprender un mapa orientarse, imaginarse la disposición de unos muebles en un espacio determinado o en la proyección de la trayectoria de un objeto móvil. Es la propia de los aviadores, exploradores y escultores.
5. **Inteligencia cinestésico-corporal:** se utiliza en la ejecución de los deportes, de bailes y en general en aquellas actividades donde el control corporal es esencial para obtener un buen rendimiento, propio de bailarines, gimnastas y mimos.
6. **Inteligencia interpersonal:** se aplica en la relación con otras personas para comprender sus motivos, deseos, emociones y comportamientos. Es la capacidad de entender y comprender los estados de ánimo de los otros, las motivaciones o los estados psicológicos de los demás. Se refiere a la capacidad cognitiva de comprender. Se encuentra muy desarrollada en maestros, vendedores o terapeutas.
7. **Inteligencia intrapersonal:** la capacidad de acceder a las sentimientos propios, las emociones de uno mismo y utilizarlos para guiar el comportamiento y la conducta del mismo sujeto, se refiere a una capacidad cognitiva de comprender los estados de ánimo de uno mismo. Juega un papel determinante en los cambios personales asociados a mejoras o adaptaciones a los eventos vitales. Se encuentra en monjes, religiosos y yoguis (Pueyo, 1999 p. 6).

De acuerdo con Gardner esta lista representa en su conjunto los objetivos valorados en diferentes culturas, sobre la dotación de estas inteligencias que le permitirán a un individuo, según su diferente dotación ser capaz de resolver problemas o diseñar productos. La teoría de la Inteligencia Múltiple expuesta por el autor representa un esfuerzo (en palabras del mismo) "un esfuerzo por fundamentar de forma amplia el concepto de inteligencia en los más amplios conocimientos científicos actuales posibles, pretende ofrecer un conjunto de herramientas a los educadores con las que ayudar al desarrollo de las potencialidades individuales, y creo que aplicada de forma adecuada puede ayudar a que todos los individuos alcancen el máximo desarrollo de su potencial tanto en la vida profesional como privada" (Gardner en Pueyo, 1999. p. 7)

### **2.3.3 Teorías sobre la inteligencia emocional.**

La inteligencia emocional es definida "como la habilidad para percibir, expresar y valorar con exactitud las emociones; como la habilidad para generar sentimientos que faciliten el pensamiento; como la habilidad para entender las emociones y el conocimiento emocional y, finalmente, como la habilidad para regular para regular, reflexivamente, las conductas emocionales de tal manera que favorezcan el crecimiento intelectual y emocional (...). Esta definición enfatiza que la emoción hace pensar más inteligentemente y que uno puede pensar inteligentemente sobre la vida emocional" (Acevedo y Acevedo, 1998. p.44)

La inteligencia emocional se sustenta en: la amplitud de la emotividad personal, cuanto más variadas sean las emociones que experimente el sujeto más riqueza de pensamientos evocará sobre ellas. Goleman (1995 en Acevedo y Acevedo, 1998) plantea que el modelo se sustenta en datos empíricos de la realidad social, expresados a través de los efectos que producen conductas como la agresividad, la ansiedad, la impulsividad, etc., es decir, en la dimensión emocional y la forma en que dicha expresión moviliza el funcionamiento intelectual. Considera que el constructo tiene que ver más "con la capacidad de motivarse así mismo, con las expectativas que se poseen, con la persistencia en las frustraciones, con la autorregulación de los impulsos y del saber esperar, que con los índices académicos o profesionales que obtenga la persona - Investigaciones realizadas por el autor lo llevaron a concluir que - la inteligencia académica no nos prepara para la vicisitudes que nos depara la vida. Considera

que la escuela debe ejercer una función más directa en la formación de las actitudes de los sentimientos y del autoconcepto de los niños (p. 44-45)

Por su parte Gardner (1993) y Sternberg (1986) reconocen la importancia que tienen las emociones para el primer autor en su teoría de las inteligencias múltiples en la inteligencia que denomina interpersonal e intrapersonal y Sternberg en la inteligencia práctica o inteligencia social. Para resumir estos aspectos teóricos de la inteligencia emocional podemos decir que está siendo muy investigada en la actualidad como modelo independiente y en interacción con otras explicaciones más tradicionales de la inteligencia, como la perspectiva cognitiva

Para Finalizar este apartado retomaremos el análisis que realiza Mora (1993) presentando un amplio resumen sobre las fortalezas y debilidades que presentan los diversos modelos explicativos de la inteligencia, algunos de ellos abordados en los apartados anteriores y otros no expuestos, pero que son considerados por el autor y que para tener una visión más completa del tema son también considerados:

#### **2.3.4 Fortalezas y debilidades de los diversos modelos explicativos de la inteligencia.**

**Primer modelo: geográfico;** centra su atención en la forma que toma el mapa de la mente, quizá como fruto inmediato de una visión del cerebro como un conjunto de capacidades perfectamente 'localizadas', como era lo habitual en las primeras décadas del siglo XX. Dentro de este contexto se elaboran teorías típicas como los dos factores cruciales de la inteligencia "g" y "s" (Spearman, 1904, 1923) o las habilidades mentales primarias (Thurstone, 1972) o los primeros modelos de la

estructura del intelecto de Guilford (1967). El factor se presenta como la unidad típica de análisis y como metodología básica el análisis factorial. Entre sus características más positivas, se encuentran unas nítidas estructuras mentales directamente operacionalizables a través de los test mentales; y el desarrollo de una sofisticada y adecuada máquina estadística, el análisis factorial, para manipular los resultados en la ejecución de los mismos, sin embargo, al poner el énfasis en el mapa de la mente, deja de lado el proceso mental como tal.

**Segundo modelo: computacional**, el énfasis explicativo se centra en las rutinas, (programas) sobre entendidas en lo que denominamos 'pensamiento inteligente', algunos ejemplos de este modelo serían: la velocidad mental de Jensen (1973), la eficacia verbal de Hunt o la subteoría componencial de R.J. Sternberg (1982), en estas teorías a la hora de explicarnos qué es la inteligencia se van a elegir procesos de información elementales como por ejemplo el análisis de tiempo de reacción, o se van a construir programas de simulación en ordenador, como metodología típica de aproximación a la conducta inteligente, la ventaja que presenta este modelo es que nos permite obtener una especificación detallada de los procesos y las estrategias mentales, analizar con precisión variables y obtener el tiempo de realización de la tarea, sin embargo presenta limitaciones en cuanto a que la mente humana no puede ser analizada como un programa de ordenador, y que sus cuestiones no se pueden generalizar a la inteligencia de "cada día".

**Tercer. Modelo: antropológico**, representado por autores como Berry, Cole o Charlesworth, quienes "tienen en común el análisis del contexto cultural, como el camino más idóneo para definir qué entendemos por inteligencia, aunque no coincidan totalmente en que deba sobreentenderse por este concepto, (...) en

Berry puede ser definida como un relativismo cultural radical; la de Cole como un Comparativismo condicional, mientras que la de Charlesworth sería una descripción del marco ecológico” (Barry, en Mora, 1993, p. 288) Su principal aportación vendría dada por el reconocimiento del rol sociocultural en la determinación de lo que se entiende por conducta inteligente, consigue aterrizar la teoría en el contexto de la vida cotidiana, pero presenta algunas imprecisiones, como son: falta de concreción, de definiciones culturales, de falta de parsimonia científica, la falta de poder concretar aspectos tan cruciales como el funcionamiento cognoscitivo de los sujetos en el marco en que se está analizando

**Cuarto. Modelo: Psicogenético.** Desarrollado por Piaget con su epistemología genética, que aunque su objetivo no es construir una teoría de la inteligencia, ésta se encuentra implícita, ya que plantea cómo se desarrolla en cuanto sistema, tanto filogenética como ontogenéticamente. Las principales fortalezas de modelo piagetano han sido el reconocimiento de la importancia del desarrollo y la amplitud de la teoría para incorporar en su seno al pensamiento científico y a la génesis del desarrollo de la infancia. Las críticas que se le suelen hacer son la sobreestimación de las edades mínimas para conseguir las competencias cognitivas, la no referencia a formas de pensamiento ajenas a las no-científicas o no-lógicas, la propia cuestionabilidad del concepto de fase de desarrollo y cuál debe ser la tabla de éstos a la que adherimos.

**Quinto Modelo: Sociológico.** Plantea que la inteligencia es algo relativo a la experiencia en el marco social del individuo. Un ejemplo de éste es la teoría sobre la zona de desarrollo próximo de Vygotsky (1962 y 1978) o las experiencias de aprendizaje mediatizado explicitadas por Fuerstein (1980) “en ambas teorías se

intenta reconocer lo que denominamos inteligencia a través del aprendizaje mediatizado por la experiencia, fundamentalmente en niños" ( ). Su metodología se basa en lo que se conoce como entrenamiento cognoscitivo y su posibilidad de mejora. El principal logro de este modelo para entender la inteligencia viene representado por el reconocimiento de la internalización de las experiencias iniciales y el papel mediador que ejercen en la génesis y el desarrollo de todo lo que denominamos destrezas cognoscitivas, tanto de los padres como del contexto social, en el que se mueve el niño" (Mora, 1993, p. 288) La debilidad del modelo se presenta en el especificar el cómo y el cuándo del proceso de internalización y la falta de instrumentos fiables que midan los conceptos manejados en dicha teoría.

Finalmente se presenta **el modelo del autogobierno mental** de la inteligencia, de Sternberg, que vendría a ser un resumen de los anteriores, tratando de superar sus limitaciones, bajo este modelo se busca establecer el cómo los individuos se rigen y se gobiernan así mismos. "En dicho modelo se van a mezclar aspectos del mundo externo, aspectos del mundo interno y aspectos de la experiencia del individuo ( ). Este modelo ha sido denominado también teoría triárquica de la cual se desprenden tres subteorías: Componencial (pensamiento analítico); experiencial (pensamiento creativo) y el contextual (cómo aprendemos a manipular nuestro entorno, el conjunto de acciones personales y sociales) (Sternberg, en Mora, 1993) El modelo puede ser considerado como uno de los más completos para explicar lo que es la inteligencia.

En términos generales podemos decir que aunque en las diferentes teorías expuestas existen diferencias en cuanto a sus niveles de elaboración o

sistematización lo cierto es que ambas vertientes (las teorías clásicas y las teorías múltiples) buscan explicar la inteligencia en función de componentes integrantes de la elaboración de la información y para demostrarlo se basan principalmente en el manejo de ejercicios los cuales por lo general son estudiados a través de los llamados test de inteligencia, lo que nos remite a nuestro tercer apartado

### **2.3.5 Los test de inteligencia; origen, tipos y características**

Para iniciar este apartado se considera conveniente definir en primera instancia lo que es un test.

#### **2.3.5.1 Definiciones:**

La palabra inglesa test tiene varios significados "como nombre significa prueba, reactivo, etc; como verbo ensayar, probar o comprobar. Los lingüistas llevan su origen al latín 'testis', raíz que figura en numerosas palabras castellanas, como testigo, testimonio, etc " (Arias, 1995 P 30-31).

Anastasi (1980) considera que un test psicológico constituye esencialmente una medida objetiva, y tipificada de una muestra de conducta (p.21)

Morales (2000) sintetiza una definición en la que a partir de los factores comunes encontrados enuncia que "un instrumento de medición psicológica es aquella técnica metodológica producida artificialmente, que obedece a reglas explícitas y coloca al o a los individuos en condiciones experimentales, con el fin de extraer el segmento del comportamiento a estudiar y que permitan la comparación estadística conductora a la clasificación cualitativa, topológica o cuantitativa de la característica o características que se están evaluando" (p.17)

Pieron considera que el test mental "es una prueba definida, implicando una tarea a realizar, idéntica para todos los sujetos examinados. El test debe disponer de una técnica precisa mediante la cual se pueda apreciar el éxito o el fracaso del examinado, o bien de una noticia numérica aplicable al resultado. La tarea puede consistir en poner de manifiesto conocimientos adquiridos -test pedagógicos-, o bien funciones sensoriomotrices o mentales -test psicológicos-" (en Cerda, 1973 p.48)

Por su parte Crombach (Morales, 2000) considera que una prueba psicológica es una técnica sistemática que compara la conducta de dos o mas personas

En general se puede considerar el test "como un reactivo que, aplicado a un sujeto, revela y da testimonio del tipo o grado de aptitud, de su forma de ser o del grado de instrucción que posee. Estos reactivos o test constan de preguntas, tareas, estímulos, situaciones, etc. Que intentan poner de relieve una muestra de las conductas del sujeto, representativas de las características que se quieren apreciar o medir" (Martínez. 1995, p 30-31).

Como se puede observar, todas estas definiciones coinciden con el hecho de someter al sujeto de observación a una situación de prueba con la finalidad de conocer su comportamiento, ya sea éste psicológico o pedagógico básicamente, y poder con ello establecer desde criterios de tipo normativo, hasta aquellas cualidades inherentes a su persona. Por tanto, los objetivos que se persiguen a través del uso de este tipo de instrumentos va a depender de los intereses que persigue quien los utiliza. Para conocer un poco más acerca de estos valiosos instrumentos en el siguiente punto se expondrá cuál es su origen y los principales experimentos que se llevaron a cabo, para llegar es establecer su confiabilidad y su validez como instrumentos de medición científica, especialmente de los test que valoran capacidades intelectuales o más comúnmente conocidos como test de inteligencia

### ***2.3.5.2 Origen de los test de inteligencia***

El origen de los primeros test se pierde en la antigüedad, en algunos estudios realizados se ha localizado información en la que se puede inferir el uso de exámenes para estimar el dominio tanto físico como mental de los sujetos, en el antiguo Imperio chino se tienen antecedentes de la aplicación de exámenes para la administración pública, entre los griegos, constituían un complemento integrado en

el sistema educativo el método socrático de enseñanza es otro ejemplo, con su entretendido de preguntas y respuestas, de lo que hoy conocemos como enseñanza programada. en la Edad Media en Europa, las universidades se basaron en los resultados obtenidos en exámenes regulares para conceder grados y honores (Anastasi, 1970). Lo expuesto nos lleva a pensar que el hombre desde la antigüedad a estado preocupado por establecer mecanismos que le permitan, valorar a los sujetos en función de sus potencialidades, o sus deficiencias, y es precisamente a partir de la evaluación de éstas últimas donde los test tal y como los conocemos hoy, adquirieron relevancia.

La idea de obtener instrumentos efectivos que ayudaran a la identificación y clasificación de los débiles mentales llevó a los primeros psicólogos experimentales del siglo XIX, a realizar una serie de pruebas que les permitieran disminuir el grado error que se tenía al realizar sus observaciones, ya que el objetivo principal de sus estudios giraba en torno a la formulación de descripciones generalizadas de la conducta humana, quedando las diferencias individuales como males necesarios para la aplicación de las generalizaciones, por lo que la presencia de tal error, o variabilidad individual, hacia que las generalizaciones fueran solo aproximadas. Sin embargo las observaciones realizadas los fue llevando a un cambio de actitud frente a tales diferencias, centrando su atención en los fenómenos sensoriales y en el control riguroso de las condiciones en las que se hacia la observación de todos los sujetos bajo procedimientos tipificados; y son precisamente éstos, los que le dieron a los test psicológicos una de sus principales características. Pero veamos como es que se empezaron a construir estos instrumentos de medición en los estudios de laboratorio.

### *Construcción de los primeros test mentales:*

Se considera a Sir Francis Galton de Inglaterra una de las figuras claves en el desarrollo de la medición mental, ya que intentó demostrar a través de la aplicación de una serie de métodos las bases hereditarias de la inteligencia. (Zambrano, 1998) Galton creía que los test de discriminación sensorial podían servir para medir el intelecto, idea que le fue sugerida por las teorías de Locke, en sus reflexiones consideraba que "la única información que llega hasta nosotros sobre los acontecimientos exteriores pasa, al parecer, por la avenida de nuestros sentidos: y cuanto más perciban la diferencia nuestros sentidos, mayor será el campo sobre el que puede actuar nuestro juicio e inteligencia". (Galton, en Anastasi, 1980. p.8) Para tratar de demostrar con bases científicas sus planteamientos Galton fue uno de los primeros en metrizar cualquier fenómeno social que estuviera a su alcance, incluyendo por supuesto los mentales, psíquicos y comportamentales, por lo que se crearon diversos instrumentos psicológicos, entre los que se encuentra la conocida balanza de presión para la medida de los umbrales, construida por Wundt, que influyó de manera directa en el surgimiento de los test mentales, principalmente en los test de inteligencia.

Las investigaciones realizadas sobre el cálculo de los umbrales sensoriales, y la influencia que éstos tuvieron en la creación de las pruebas de inteligencia se debe a que a través de ellos se pretende determinar la importancia que tienen las técnicas, las tradiciones operativas, los instrumentos, etc., en el desarrollo de las teorías científicas, es decir, si queremos explicar lo que es la inteligencia en términos científicos, debemos entender el papel que juegan los instrumentos antes mencionados para la construcción de éstas

Sobre este aspecto encontramos los trabajos desarrollados por Solla Price (1984), quien es uno de los principales fundadores de los análisis cuantitativos sobre el fenómeno científico, la principal tesis de este autor se centra en lo que él ha denominado "instrumentalidad: 'un método de laboratorio para hacer algo a la naturaleza o a los datos disponibles'. En algunos casos, las instrumentalidades son técnicas, en otros son instrumentos, y algunas veces son procedimientos" (Lujan, 1990. p 42). Bajo este concepto el autor explica que la ciencia no proporciona a la tecnología teorías para ser aplicadas, sino precisamente instrumentalidades.

Lujan (1991) parte de los planteamientos de Price para explicar la relación que guarda la psicología científica, los laboratorios psicológicos y el empleo de los instrumentos psicofisiológicos, para definir la importancia que tiene el contexto técnico -instrumental (definido por él mismo) en el desarrollo de las pruebas de inteligencia.

Los primeros instrumentos para medir procesos mentales fueron el cronoscopio y el crónografo. Los cronoscopios son simplemente mecanismos de relojería de gran precisión que permiten medir el tiempo que un individuo tarda en dar una respuesta tras haber recibido un estímulo. El crónografo era un instrumento con dos agujas, una de las cuales trazaba una línea cuando el sujeto recibía un estímulo mientras otra lo hacía cuando efectuaba la respuesta. De este modo se podía medir el tiempo de reacción comparando las dos líneas trazadas por el crónografo. (Lujan, 1990 )

Continuando con las aportaciones que los primeros psicólogos experimentales realizaron para la elaboración de los test de inteligencia desarrollaremos los trabajos realizados por Galton, Cattell, Binet Terman, entre otros sobre la Inteligencia. En lo que se refiere a las investigaciones desarrolladas por Galton se encuentran las relacionadas a la fundamentación científica de la eugeneia, de donde nace la teoría hereditaria de la inteligencia y la eugeneia, teniendo como hipótesis central "demostrar que la elevada reputación (o eminencia) es una prueba bastante fiel de la elevada aptitud". - Para comprobar dicha hipótesis Galton recurre a los trabajos realizados por Lambert, Adolphe y Quetelet sobre las distribución de los rasgos físicos, aplicando la ley de la desviación de la media, conocida actualmente como la ley de distribución normal a sus estudios, llevándolo a la afirmación siguiente:- "si la reputación o reconocimiento público son una buena medida de aptitud natural, entonces deben distribirse de acuerdo con la ley de la desviación de la media, la ley que gobierna las desviaciones de todas las medidas verdaderas" -Para demostrar lo anterior-" Galton calculó con datos reales, que 250 individuos por millón llegan al grado más alto de eminencia. Proporción que coincidía, asombrosamente, con sus cálculos teóricos de la frecuencia de los individuos de la clase X (la de los individuos con mayor aptitud natural): 248 por millón. Concluyo entonces que la eminencia (o reputación pública) era una buena medida de aptitud natural"(Lujan, 1990. p. 44-45) No conforme con estos resultados Galton reconoce que también la medida de los tiempos de reacción podían de alguna manera ser considerados como una medida de la inteligencia, creando con ello toda una tradición de investigación en psicología diferencial. Tradición que será seguida hasta desaparecer por Binet y Simón, para ser retomada en la actualidad por Eysenck (1985) y Jensen (1987)

El Manejo de estos conceptos le valieron al autor una contribución más, que es el desarrollo de métodos estadísticos para el análisis de los datos sobre las diferencias individuales. Galton selecciono y adoptó cierto número de técnicas, elaboradas previamente por los matemáticos, dispuso que estas técnicas en forma tal que pudieran ser utilizados por investigadores no formados en el área de las matemáticas y que desearan manejar cuantitativamente los resultados de los test, lo que permitió que se extendiera enormemente la aplicación de los procedimientos estadísticos a los análisis de los datos de los test. (Anastasi, 1988)

Los trabajos de Galton fueron también seguidos por otro eugenista convencido, James Mckeen Cattell, discípulo de Wunt. En un artículo escrito por el autor en 1890, el autor utiliza por primera vez el término "test mental" para designar una serie de pruebas psicológicas, con la finalidad de determinar el nivel intelectual de un grupo de universitarios "La mayor parte de estos test estaban destinados a medir funciones sensoriomotrices tales como fuerza muscular, tiempo de reacción, sensibilidad al dolor, discriminación de pesos, rapidez de movimientos, etc.," (Cerdeña, 1973, p.49) pruebas que permitieran demostrar y apoyar la teoría galtoniana de la aptitud mental. Sin embargo estas argumentaciones pronto fueron echadas por tierra, por un discípulo de Cattell llamado Clark Wissler, quien hizo evidente que los resultados obtenidos en los test mentales, basados en las reacciones fisiológicas, poco se correlacionaban con las puntuaciones académicas obtenidas por 300 alumnos de la Columbia University y el Barnard College, lo que lleva finalmente a que por un tiempo se perdiera el interés por este tipo de trabajos.

Ya para entonces, Binet y Víctor Henri en 1885 critican en el trabajo titulado la "Psicología individual" los resultados obtenidos por los test utilizados hasta el momento considerando que eran demasiado sensoriales y concentrados exclusivamente en el estudio de las funciones simples por lo que proponen nuevas series de test adecuados para explorar las funciones más complejas, como la memoria, imaginación, atención, comprensión, sugestibilidad, apreciaciones estéticas, y muchas otras. En estos test se pueden identificar muy fácilmente las tendencias para el desarrollo de las famosas escalas de inteligencia, de Binet. (Anastasi, 1980; Cerda, 1976).

En 1896, dieron un giro fundamental en la forma de entender los test mentales, ubicándolos dentro de lo que ambos consideraron como un nuevo campo de investigación, que en gran parte redefinió la labor y el objeto de la psicología general, es decir, la creación de la **psicología individual**: "la psicología general estudia las propiedades generales de los procesos psicológicos, y por lo tanto son comunes a todos los individuos; la psicología individual, por el contrario, estudia aquellas propiedades de los procesos psicológicos que varían de un individuo a otro" (Binet en Lujan, 1991, p.47). Haciendo también uso de la medición de los umbrales sensoriales para la elaboración de sus test de inteligencia, Binet y Simón partieron de una definición operacional para calcular umbrales intelectuales en un contexto social determinado del que depende el grado de dificultad que esta en consonancia con su concepción de inteligencia como un ejercicio en la vida cotidiana de múltiples facultades psicológicas enlazadas bajo el control del razonamiento práctico. Estos razonamientos llevaron a los autores a referir el núcleo de la inteligencia en términos de juicio, también llamado sentido común,

iniciativa, facultad de adaptarse a las circunstancias. Juzgar bien, comprender bien, razonar bien, como actividades esenciales de la inteligencia

Binet sugirió que el pensamiento tiene tres elementos distintos a los que llamó *dirección*: que implica saber qué hay que hacer y cómo hacerlo; *adaptación*: se refiere a la creación de una estrategia para realizar una tarea, luego conservar la huella de la estrategia y adaptarla al mismo tiempo que se la aplica; y *crítica* que es la habilidad para criticar los pensamientos y las acciones propias (Sternberg, 1997, p.59).

Lo que estos autores pretendían era determinar las diferencias individuales, a partir de procesos más complejos e interactivos, tales como las facultades psíquicas superiores y no por los procesos y sensaciones más elementales. En la primera escala desarrollada se incluyeron 30 problemas o test, colocados en orden creciente de dificultad, éste fue determinado de forma empírica aplicando los test a 50 niños normales entre los 3 y los 11 años, y algunos niños retrasados mentales, el objetivo de los test era cubrir una amplia variedad de funciones, con especial preferencia para el juicio, la comprensión y el razonamiento, que Binet consideraba componentes esenciales de la inteligencia. En la segunda versión de esta prueba se agruparon los test en niveles de edad, de manera que las puntuaciones obtenidas por un niño en el test pudiera expresar como *edad mental*, concepto que permitió en gran medida que se popularizara la aplicación de los test de inteligencia (Anastasi, 1980).

Binet dio un fuerte impulso a los test de inteligencia y los colocó en un lugar que les permitió producir resultados fructíferos. Un ejemplo de ello son los trabajos

desarrollados por él y Simon, sobre tres posibles métodos para medir la inteligencia:

- “1. El método médico, que trata de apreciar los signos anatómicos, fisiológicos y patológicos de la inteligencia inferior
2. El método pedagógico, que trata de juzgar la inteligencia de acuerdo con la suma del conocimiento adquirido.
- 3 El método psicológico, que realiza observaciones y medidas directas del grado de inteligencia” (Lujan, 1991. p 49)

Los dos primeros métodos presentaban limitaciones para determinar el nivel de inteligencia, por lo que optaron por desarrollar el tercero a través de la reunión de una serie de tareas que no se quedaran en definir lo mental a partir de lo físico, como el primer método, o a partir del nivel de instrucción como lo implicaba el segundo, sino que la idea fundamental de este método significaba el establecimiento de lo que llamarían una **escala de inteligencia** la cual estaba compuesta por una serie de test de dificultad creciente, empezando por el nivel intelectual más bajo que puede observarse, y terminando en la inteligencia normal media. Cada grupo en la serie corresponde a un nivel mental diferente. Ordenaron su escala según las edades de los niños con los que trabajaron, a partir del criterio de que si la mayoría de los niños de un determinado grupo pasaban el test, éste era entonces el nivel de edad correspondiente de dicho test

Los test de Binet causaron un gran impacto en los Estados Unidos, en 1916 Lous Terman, de la Universidad de Stanford, realizó una adaptación del test, la cual se conoce con el nombre de Test de Inteligencia Stanford-Binet, a través de

éste se dio origen a un nuevo concepto: *el cociente intelectual* o *CI* o razón entre la edad mental y la edad cronológica real, por 100 para eliminar decimales. Este test no tardo en convertirse en un instrumento estándar para la psicología clínica, la psiquiatría, y el asesoramiento educativo. A partir de los resultados obtenidos con la aplicación del test. Terman y otros empezaron a creer a diferencia de Binet, que la inteligencia era una cualidad interna, que tenía un fuerte componente hereditario, y que por tanto los test del CI podían medir esta desigualdad (Zambrano, 1988)

La diferencia fundamental entre los trabajos revisados radica en que Binet no se preocupó por el tema de la herencia de la conducta, sino que centró sus esfuerzos en el desarrollo de programas de aplicación práctica, es decir una "ortopedia mental" (que en terminología actual debería denominarse, según el autor 'educación compensatoria') destinada a resolver la problemática que representaba la instrucción de los niños retardados. Para él la inteligencia dependía del juicio mental y no de la agudeza sensorial; lo que lleva a la siguiente reflexión: las posturas de los autores revisados nos lleva hacer referencia a la polémica "herencia versus ambiente en el seno de la psicometría. Lo que Galton pretendía medir para fundamentar su programa eugenésico era lo que se podría denominar inteligencia innata, es decir, la inteligencia de origen biológico" (tesis sostenida posteriormente por Terman). -Lo que Binet en un principio pretendió medir fue- "la inteligencia de origen social, utilizó el concepto de inteligencia como constructo de consenso social con el fin de reivindicar su programa de educación compensatoria", -que posteriormente fue retomado a su vez por el movimiento eugenista, lo que produjo una- "situación de gran trascendencia en las investigaciones sobre inteligencia: la identificación entre inteligencia social e inteligencia biológica", (...) a partir de este momento se considerará que las diferencias en inteligencia social son de origen biológico". (Lujan, 1991, p. 49-50)

Otro autor que dio un gran impulso a uso de los test de inteligencia, fue el trabajo realizado por David Wechsler en 1939, quien confeccionó un nuevo test con el fin de resolver el mayor problema que existe en general con los test de inteligencia: su dependencia del lenguaje, un problema para las personas que no hablan el idioma de origen de la prueba y para los niños pequeños, por lo que

desarrolló nuevas secciones no verbales paralelas a las verbales. (Zambrano, 1998) Además de que por primera vez se llevo a cabo la estandarización de un test de inteligencia para adultos en función de su edad cronológica. Compuesta en parte por escalas tipo Binet y Yerkes-Brdges y Hardwick representan el punto de partida de numerosos test de aplicación colectiva (Cerde, 1973)

Como se puede ver la construcción de los test inteligencia paso por diversos momentos, llevando una y otra vez a los investigadores a la confirmación de sus hipótesis y al planteamiento de otras nuevas, generando con ello la elaboración de nuevos instrumentos que fueran más válidos y confiables. como es el caso de las construcción de los test de factor "g", veamos brevemente en que consisten

### **2.3.5.3 El factor "g".**

De acuerdo con Carroll (1998), hay pruebas contundentes de la existencia de un factor "general" de inteligencia que se encuentra implicado en gran cantidad de tareas cognitivas. Es probable que éste sea el mismo factor que Charles Spearman llamó g en 1904, quien proponía la existencia de un factor de inteligencia general que representaba lo que es común a las habilidades que constituyen el Coeficiente Intelectual IQ. Según esta teoría todas las ramas de la actividad intelectual tienen en común una función intelectual, ese factor g, que permite que las personas se manejen con la misma facilidad en distintas tareas las cuales implican la 'educación de correlatos y de relaciones'. Siempre que una tarea requiera descubrir semejanzas y deferencias entre elementos, inferir correspondencias, reglas y generalizaciones, seguir una línea de razonamiento y

predecir consecuencias, esa tarea probablemente implica este factor general de inteligencia especialmente conforme los elementos de la tarea son más numerosos y complejos (Sanz, 2001; Carrol, 1988) Este planteamiento no solo se limita a la resolución de problemas expuestos en los test psicológicos, sino también en otras tareas tanto académicas, como técnicas, prácticas y sociales.

Por otra parte el factor "g" , sugiere que la naturaleza de las diferencias individuales puede ser entendida en términos de los niveles de energía mental que el individuo utiliza en la realización de una tarea intelectual, y/o en términos de aptitudes que el sujeto utiliza para realizar la tarea a partir de tres principios cualitativos: percepción de la experiencia, educación de relaciones y educación de elementos correlacionados Sobre el primer aspecto (niveles de energía) se encuentra un estudio realizado por Duncan, a través del cual trató de determinar científicamente si el factor g reside en un área cerebral concreta o si por el contrario, está más ampliamente distribuido en el cerebro.

Para demostrar lo anterior el Dr Duncan junto con un equipo de investigadores británicos y alemanes, sometieron a un grupo de voluntarios a una serie de tareas espaciales, verbales, perceptivas y motrices con una alta implicación del factor g, en paralelo a tareas de control con baja implicación de este factor, con las que se contrastaban. Simultáneamente al desarrollo de las tareas, los cerebros de los voluntarios se escasearon empleando la Tomografía de Emisión de Positrones, que es una técnica de exploración por neuroimagen que permite estudiar , *in vivo*, los aspectos estructurales y funcionales del sistema nervioso central. Mediante el experimento, los investigadores observaron que las tareas en las que se implicaba un trabajo de g-alta, no requería la intervención de

numerosas regiones del cerebro de forma conjunta, sino que la actividad se centraba en una zona; el cortex lateral, en uno o ambos hemisferios cerebrales. Los resultados sugirieron que la inteligencia general proviene de un sistema frontal específico importante en el control de distintas formas de comportamiento, por lo que se puede concluir a decir de los investigadores que el estudio proporciona una base científica para la localización física de la "morada" de la inteligencia general (Sanz, 2001)

En cuanto a las aptitudes que el sujeto utiliza, Spearman (en Cerda, 1973) aplicando el método de concordancias y diferencias a una muy considerable serie de matrices de correlación y principalmente a los resultados obtenidos en un experimento a gran escala que se llevó a cabo en Chicago, el autor concluyó que la "nota característica de g es la neogénesis abstracta. El factor g es considerado como una propiedad común a todas las operaciones cognoscitivas, y se manifiesta en alto grado en las funciones abstractas y conducentes a nuevos conocimientos por medio de la educación de relaciones y correlatos (...) Este factor g satura altamente test de abstracción, de razonamiento y de aprehensión de relaciones y correlatos (p. 131-132).

Por su parte para Jensen (1992) el análisis factorial proporciona una estimación cuantitativa del grado relativo en que la varianza de cada uno de los test de un grupo determinado, está "saturado" por la fuente de varianza que es común a todos los test del conjunto, esta fuente común es a la que el autor denomina factor general o simplemente g, que en un amplio conjunto de diversos test mentales constituye una definición operatoria de inteligencia, científica y prácticamente útil.

Ubicado dentro de este contexto encontramos los trabajos realizados por R.B Cattell (1971), quien analiza el problema de la medición de la inteligencia desde lo que él denomina inteligencia *fluida* e inteligencia *cristalizada*; "la primera representa la capacidad básica y la segunda representa las capacidades requeridas a través del aprendizaje, la práctica y la exposición a la educación" (Carroll, 1982, p. 118) . En términos más amplios la inteligencia fluida implicaría la posesión por parte del sujeto de una cantidad suficiente de inteligencia general, lo que permite que una persona adquiera los correspondientes niveles de aptitud en el lenguaje, matemáticas y otros sistemas simbólicos, lo que es posible solo si se da la suficiente exposición de éstos, a través de la educación, el entrenamiento y la experiencia; y dependiendo del grado en el que el sujeto se haya beneficiado de esta exposición generará la llamada inteligencia cristalizada. Por lo tanto, el rendimiento y el grado en que una persona se enfrenta a diversas tareas y problemas, refleja las aptitudes generales de inteligencia fluida y cristalizada

Cattell genera de estos planteamientos un modelo jerárquico de las capacidades intelectuales en donde a partir de la correlación de 20 factores primarios, revela la existencia de seis factores oblicuos de segundo orden y los cuales clasifica en: inteligencia fluida, inteligencia cristalizada, capacidad de visualización, capacidad general de visualización, velocidad general de percepción y capacidad general de memoria, demostrando a su vez la existencia de otros tres factores de tercer orden a los que definió como: inteligencia fluida original, capacidad de concentración y cultura escolar.

Sobre los modelos jerárquicos de la inteligencia Vernon (1980) presenta un modelo en el que se describe la mente como una especie de jerarquía o árbol genealógico, en el que el factor g es el componente más destacado en el sentido de que da cuenta de la proporción de las diferencias en aptitudes, las cuales tienden a caer en dos tipos principales: el verbal educativo y el espacial-perceptivo-práctico.

La construcción del factor "g" llevó a los autores mencionados y a otros investigadores (Eysenk, Anastasi en Castelló, 1995, p. 205-206), a la búsqueda de una medida más pura de inteligencia, centrando su atención en la creación de instrumentos que les permitieran eliminar los efectos culturales o aquellos ligados a la información, manteniendo los más puramente procesuales, aportando como resultado que las aptitudes vinculadas al proceso lógico, a la inducción y a la generalización, constituyeran el núcleo de los instrumentos de factor "g", depurados de elementos verbales (muy ligados al origen sociocultural) y minimizando el acceso a la memoria con el ánimo de convertirlos en medidas libres de cultura. Esta condición posiblemente es la que ha permitido que este tipo de pruebas tengan aún vigencia y sigan siendo ampliamente utilizadas, ya que constituyen un instrumento de gran utilidad para evaluar la inteligencia, pero veamos cómo se construye el factor g.

#### **2.3.5.4 Construcción del factor G.**

Los diversos modelos del factor G se incluyen dentro de las teorías monolíticas y unifactoriales de la inteligencia. El factor g se desarrolló en paralelo con los estudios relacionados con el CI y ambos utilizaron sistemas de validación concurrente. Por otra parte también comparte la conceptualización de la inteligencia como una sola variable y la validación en ámbitos académicos, sin embargo presentan diferencias en cuanto a dos aspectos fundamentales:

- a) La metodología utilizada en los modelos de factor G es el principal esqueleto teórico de los mismos, el método matemático del análisis factorial (especialmente del centroide y no rotado) aplicado a los test intelectuales de los años veinte genera la identificación de los elementos comunes a los mismos. Estos elementos comunes, aglutinados en el primer factor centroide, constituyen la base del concepto de inteligencia general (factor g), aunque esta generalidad esté restringida a lo medido por los instrumentos factorializados.
  
- b) Los efectos de la detección de aquello que los test intelectuales tenían en común, sumada a tal disposición, de eliminar los efectos culturales o ligados a la información, manteniendo aquellos más puramente procesuales, aportaron como resultado que las aptitudes vinculadas al procesamiento lógico y a la inducción y generalización constituyen el núcleo de los instrumentos de factor g en trabajos posteriores, estos test han sido depurados de elementos verbales (muy relacionado con el origen sociocultural) y de información (minimizando el acceso a la memoria) con el ánimo de convertirlos en medidas

libres de cultura (Anastasi en Castelló 1995). En estos tipos de test el razonamiento lógico y los procesos inductivos-deductivos, son los factores que más sólidamente aparecen representados. En conclusión se puede decir que los modelos de factor g son fundamentalmente equivalentes a los del CI.

### **2.3.5.5 Características de los test de Inteligencia**

De acuerdo con Martínez (1995), los test psicométricos en general presentan entre sus principales características las siguientes:

- a) **Medida objetiva**, que implica la idea de seguridad y precisión de la medida.
- b) A partir de una **muestra de conducta**: dado un rasgo o una característica del sujeto que se pretende medir, se elegirán para formar el test elementos que constituyan una muestra representativa del conjunto de aspectos de dicho rasgo o conducta. Las respuestas concretas al test representan únicamente una pequeña proporción de la población de conductas posibles.
- c) **Técnica sistemática**: situación problemática previamente dispuesta y estudiada a la que el sujeto ha de responder siguiendo ciertas instrucciones fijadas de antemano. A la misma propiedad hace referencia la noción de estandarizada; las observaciones se recogen de forma prescrita, de tal modo que las diferencias encontradas pueden atribuirse únicamente a diferencias de los sujetos en el test y no a variaciones en la forma en que se recogen las observaciones.
- d) **Comparar conductas**: de la respuesta del sujeto se estima por comparación con un grupo normativo la calidad o grado del rasgo o aspecto a medir
- e) **Predicción o inferencia**. Un test implica normalmente predicción o inferencia acerca de conductas más importantes que las observadas durante la ejecución del test. (p. 29)

Por su parte Cerda (1973) considera que los test de inteligencia presentan cualidades como instrumentos de medida y como tales se exige que tenga una serie de características que en el fondo son las mismas que se le exigen a cualquier instrumento mensurativo, Estas características son: fidelidad, validez y sensibilidad, veamos cada una de éstas.

*Fidelidad:* Un test es fiel cuando al aplicarlo dos o más veces al mismo individuo o grupo de individuos en circunstancias similares, obtenemos resultados similares. Para establecer dicha fidelidad se debe utilizar alguno de los siguientes métodos; 1) test-retest, que implica aplicar el mismo test por dos o tres ocasiones a los mismos sujetos después de un intervalo de tiempo determinado y se calculan los coeficientes de fidelidad; 2) métodos "split-half" y "odd-even". Cualquiera de estos dos métodos consiste en dividir el test en dos partes equivalentes, administrar primero una de ellas y posteriormente la otra; 3) método de "series paralelas", consiste en construir dos test o más test paralelos, es decir equivalentes. Para ello se utilizan cuestiones de naturaleza y dificultad análoga, pero de apariencia diferente. La fidelidad de un test va a ser mayor cuanto más estable es el rango psíquico que exploremos.

*Validez:* Se dice que un test es válido cuando mide aquello que trata de medir, al igual que en la fidelidad existen diferentes métodos que pueden agruparse en tres: a) pragmáticos que consisten en la adopción de criterios objetivos exteriores al test mismo, estos pueden ser: Empíricos, utilizados fundamentalmente para los test de inteligencia para niños; escalas objetivas, ateniéndose a clasificaciones escolares para niños, rendimientos profesionales en adultos; y mediante "rating scales" elaboradas, conforme a los aspectos que trate de medir el test, por un grupo de expertos en la materia en cuestión; b) Validación apriorística, parte de una hipótesis concerniente a la naturaleza de determinado rango psíquico, y sobre esta hipótesis se construye luego un test adecuado para medirla; c) Validación factorial, se basa por un lado en el coeficiente de correlación de Pearson, y por otro en el de "unidad funcional", consistiendo en esencia en un análisis algebraico de las

covariaciones cuya finalidad estriba en poner al descubierto el orden. La relación que estas covariaciones impliquen, la validación de un test con este método se da en base a datos exclusivamente experimentales.

*Sensibilidad.* Se dice que un test es sensible o que tiene buena capacidad discriminativa, cuando es capaz de medir diferencias mínimas entre los individuos o dentro de un mismo individuo en diferentes ocasiones. Un test poseerá una mayor o menor sensibilidad según contenga más o menos ítems que permitan la subsiguiente clasificación de los sujetos y su diferenciación. (Cerdeña, 1973, p. 52-57)

Por su parte Nunnally (1991; 1970) plantea que cualquier medición científica que se pretenda realizar contiene un grado de error, dicho error se produce como un sesgo sistemático o como un error aleatorio. Esta característica se presenta también en las medidas psicológicas por lo que a la hora de elaborar un test se debe considerar su confiabilidad y su validez, aspectos que deben ser cuidadosamente tratados ya que un test puede tener un alto grado de confiabilidad y no ser válido para un propósito determinado, pero veamos a que se refiere el autor con cada uno de estos conceptos.

*Confiabilidad* se refiere a la **precisión** de la medición, independientemente de lo que se mide o más aún se puede decir que una medida es confiable cuando el error de medición debido al azar es pequeño. La confiabilidad depende del grado en que estas mediciones sean repetibles por el mismo individuo con diferentes medidas de un atributo o por diferentes personas con la misma medida de un atributo, y es precisamente este grado de confiabilidad en que le va a dar su

validez al instrumento de medición, por tanto para que un test sea de gran validez debe poseer un alto grado de confiabilidad. Pero veamos que es la validez

*Validez.* En un sentido muy general se puede decir que un instrumento de medición es válido si hace aquello para lo que está concebido, por lo tanto es importante que ésta se establezca sobre la base de investigaciones empíricas y el tipo de datos para ello depende de la clase de validez. De aquí que la validez sea una cuestión de grado y la validación un proceso continuo, en resumen; no se valida un instrumento de medición sino algún empleo al que este instrumento se aplica (Nunnally, 1970, p. 159), por lo que estaríamos hablando de función y en ese sentido las medidas psicológicas cumplen tres funciones principales:

- 1) Establecer una relación funcional con una variable particular.
- 2) Representar un universo de contenido específico.
- 3) Medir rasgos psicológicos. (Nunnally, 1991, p 100)

A estos propósitos o funciones les corresponden tres tipos de validez respectivamente;

**Validez predictiva.** Se considera cuando se usa un instrumento para estimar alguna forma importante de conducta; a esto se le denomina criterio, el cual se determina cuando el test cumple con el propósito de estimar adecuadamente la conducta que se desea evaluar. Una vez obtenido el criterio, la validez de una función de predicción se determina en forma directa y con mucha facilidad, estableciendo correlaciones entre los puntajes del test predictivo con los puntajes de la variable criterio. El tamaño de la correlación es una indicación directa de la

magnitud de la validez. El término predicción en un sentido general para referirse a relaciones funcionales entre un instrumento y sucesos ocurridos, antes, durante y después de que el instrumento se aplique, por tanto los test son sólo válidos en la medida en que cumplan bien las funciones predictivas. Finalmente la validez predictiva está determinada por el grado de correspondencia entre las dos medidas que intervienen.

**Validez de contenido.** Depende de la suficiencia con que se muestrea un dominio especificado de contenido. Para confirmar la validez del instrumento elaborado, no se recurre a la correlación como es el caso de la validez predictiva, ya que el test en sí mismo es el criterio de validez. Un ejemplo de ello sería la elaboración de cualquier examen de tipo académico para valorar el rendimiento en una unidad de enseñanza específica. Para establecer la validez de contenido es necesario asegurarse de la validez mediante el plan y los procedimientos de construcción. El plan consiste en muestrear aleatoriamente a partir de un dominio de contenido específico, y el procedimiento implicaría en seleccionar la forma en que se van a presentar los reactivos, así como las instrucciones que acompañarán el desarrollo de la prueba, por tanto, la validez de la medida se juzga por el carácter del plan y por la habilidad con que se lleva a cabo, y si además la mayoría de los usuarios de la prueba coinciden en que el plan ha sido bien ejecutado, se puede decir que dicha prueba posee un alto grado de confiabilidad. Para asegurar la validez de contenido son dos los requisitos que se deben cumplir: Un conjunto de reactivos representativos y Método sensibles para construir pruebas (Nunnally, 1991, p. 104-105).

Para determinar cuales reactivos son los que realmente valoran lo que se quiere evaluar, es necesario tener un perfil detallado o anteproyecto del tipo de preguntas o problemas que se incluirán en la prueba. de la calidad del anteproyecto se evaluará la validez del contenido. Sin embargo es importante reconocer también que las medidas que se basan en la validez de contenido, por lo general implica cuestiones de valor, es decir los valores determinan la importancia de los diferentes tipos de contenido. Otro problema que se presenta en este tipo de validez, es lo complejo que resulta asegurar el uso de métodos "sensibles" para elaborar las pruebas, por lo que se deben revisar los diversos métodos y por lo menos se debe buscar un moderado nivel de consistencia interna entre los reactivos que conforman la prueba; es decir, éstos deberán tender a medir algo en común. En síntesis la validez de contenido estriba principalmente en la propiedad del contenido y en la forma en que éste se presente.

**Validez de constructo.** Fundamentada en psicología básica se busca que a través de este tipo de validez, se establezcan relaciones funcionales entre variables relevantes (que se determina ya sea por intuición o por el contenido de las teorías psicológicas, ejemplos de este tipo de variable serían, tiempo de reacción, fuerza de hábito, inteligencia, ansiedad, etc.), que no pueden ser validadas a través de las dos anteriores. La dificultad que representa validar las variables relevantes va a depender del grado de concreción o abstracción que presente. "En la medida en que una variable es abstracta y no concreta, se dice que es un constructo, en el sentido en que es algo que el científico representa en su imaginación, algo que no existe como una dimensión aislada u observable de la conducta. Este constructo representa una hipótesis (por lo general incompleta), en la que una diversidad de conductas se correlacionan entre sí en los estudios de

diferencias individuales y/o serán afectadas de la misma manera en los tratamientos experimentales" (Nunnally, 1991, p. 110) Los constructos representan la característica fundamental de toda teoría científica.

Los constructos varían ampliamente según sea el dominio de las variables observables relacionadas: grande o pequeña; y definido estricto o vagamente. Debido a que los constructos implican dominios de observaciones, lógicamente una buena medida de cualquier constructo debería obtenerse al combinar los resultados de un cierto número de tales observaciones y no al considerar individualmente cualquiera de ellos. Por lo que es importante considerar tres aspectos fundamentales en la medición de los constructos:

- "1. Especificar el dominio de observaciones.
2. Determinar hasta que punto todos o algunos de aquellos observables se correlacionan entre sí o son afectados en forma semejante por tratamientos experimentales.
3. Determinar si una, algunas o todas las medidas de tales variables actúan como si midieran el constructo." (Nunnally, 1976, p. 168)

La validación de variables relevantes a partir de la definición de constructos abstractos resulta complejo y generalmente los hombres de ciencia se mueven entre los tres aspectos antes señalados, que lo pueden dejar limitado en la observación de los dominios o que los lleve a pruebas continuas a partir de muchos estudios de diferentes medidas propuestas del constructo y que posteriormente se acumule y evalúe la información disponible. Lo que lleva a que el investigador espere como producto final de este complejo proceso que el constructo esté bien definido en términos de un conjunto observable, para lo cual existan una o varias variables que representen bien el dominio de los observables

y que eventualmente demuestre tener una fuerte relación con otros constructos de interés. (p. 169)

### 3.3.5.6 Clasificación de los test

Existen diversas clasificaciones de los test, la gran variedad y número de test que se encuentran en el mercado hace de esto una tarea necesaria, ya que no se puede hablar de un test si tener de algún modo una clasificación, para poder entender de que test se trata. Debido a que son varias las clasificaciones que se pueden encontrar en la bibliografía especializada sobre el tema se retomará para el desarrollo de este apartado, la clasificación que realiza Martínez (1995) considerando los siguientes criterios: normas generales del método, finalidad perseguida, forma de aplicación, material empleado, características a estimar tiempo de aplicación, supuestos teórico-formales y el grado de aculturación.

- 1) Por el **método**: bajo este rubro tradicionalmente los test se han dividido en *psicométricos*, los cuales presentan las siguientes características: Se evalúan las respuestas según normas cuantitativas; todos sus elementos se valoran de forma numérica e independientemente; el resultado final es una puntuación cuantitativa; suelen medir rasgos diferenciados del sujeto *Proyectivos*, para su evaluación se siguen normas o criterios globales y cualitativos; suelen referirse a características generales y globales de la personalidad.
- 2) Por su **finalidad**: se consideran principalmente dos los criterios para establecerla según el uso que se le vaya a dar al test, éstas son: de *investigación*, si el objetivo de la aplicación del test es obtener medidas,

analizar diferencias individuales, calcular estadísticos y estimadores, probar hipótesis, etc.; de *diagnóstico* con su uso se pretende evaluar la calidad o grado de algún rasgo o de la conducta de algún individuo concreto

- 3) Por el **planteamiento del problema**: los test están contruidos con una muestra de conductas que exigen de un planteamiento del problema, cuya respuesta puede interpretarse como: *de ejecución máxima*; exige del sujeto poner en funcionamiento su capacidad a un alto grado, por lo general se encuentran dentro de este rango los test de rendimiento, aptitudes, inteligencia, entre otros. *De ejecución típica*; los elementos plantean situaciones habituales de la vida corriente, la muestra es una muestra del comportamiento típico o más frecuente del sujeto en situaciones planeadas, dentro de estos de encuentran los test de aptitudes, intereses y personalidad.
- 4) **Área del comportamiento acotada**: se distinguen los siguientes tipos; *rendimiento* académico o profesional, los test que se utilizan para evaluar esta área son de ejecución máxima destinados a medir adquisiciones de aprendizaje o conocimientos específicos, con finalidades de diagnóstico y/o pronóstico concretas; *inteligencia y aptitudes*: se refieren a rasgos psicológicos de naturaleza cognitiva (inteligencia, aptitudes específicas, creatividad, etc.) suelen ser de ejecución máxima; *personalidad, actitudes, intereses*: son test de ejecución típica con preguntas sobre la incidencia del elemento en la conducta o forma de pensar del sujeto, no en términos cognitivos, sino aspectos como forma de ser, actitudes, etc
- 5) **Por la modalidad de aplicación**: *orales*: las preguntas y respuestas son orales se usan generalmente con niños pequeños, analfabetos y personas de otras culturas: *de papel y lápiz*: se encuentra una variedad de formatos de presentación, los procedimientos de aplicación, instrucción y realización se

encuentran especificados en protocolos, cuadernillos y hojas de respuestas de forma escrita; de *manipulación o de ejecución*: en estos test al sujeto se le presentan materiales a través de los cuales se les presenta un problema a resolver. que generalmente es de coordinación psicomotriz, visualización, recortado, plegado, cubos, herramientas, etc.; *situacionales o simulaciones*: se basa en modelos de situaciones reales, para poner de relieve conductas sociales habituales; *computarizados*: la presentación del material del test y las respuestas se realiza por medio de computadoras, las cuales pueden seleccionar los ítems que debe contestar el sujeto, por esta característica se les considera además test adaptativos, y finalmente los *objetivos*: se atribuye a instrumentos específicos con registro automático de las respuestas, por ejemplo mediciones fisiológicas, tiempos de reacción, etc.

- 6) **Demandas exigidas al sujeto:** *velocidad o rapidez*; en estas pruebas los tiempos de ejecución se encuentran rigurosamente controlados y los elementos no suelen tener dificultad objetiva para los sujetos, dependiendo las diferencias individuales fundamentalmente de la rapidez de la respuesta; *potencia o dificultad*: en estas pruebas el control del tiempo no suele ser un determinante clave en el resultado. Las diferencias individuales son función de la calidad de la ejecución.
- 7) **Grado de aculturación del test:** dentro de estos se encuentran dos extremos, los más culturales en los que se encuentran manifestaciones de conducta o ejecución de claro contenido cultural, por ejemplo. los test verbales; en el otro extremo los menos influidos por la cultura para su resolución , dentro de estos se encuentran los perceptivos y visuales, en los que la experiencia cultural anterior de los sujetos tiene escasa influencia.

- 8) **Los sujetos a los que va dirigido el test.** Por edades o por grupos especiales.
- 9) **Por el modelo estadístico en el que se basa:** la construcción de cualquier test debe fundamentarse en algún modelo que permita explicar el significado de las puntuaciones y hacer inferencias a partir de las muestras de conducta planteadas en el test. Son tres los principales modelos en los que se basa la construcción de los test: *teoría clásica*, basada en el Modelo Lineal de Regresión con dos variables; *teoría de la generalizabilidad*, basada en el Modelo Lineal del Diseño Experimental y *teoría de la Respuesta al Item*, basada en modelizar las probabilidades de respuesta correcta a un ítem según las funciones de distribución normal y logística.
- 10) **Según del modelo conceptual:** dentro de los modelos conceptuales de los atributos humanos que intenta medir el test, se localizan principalmente tres concepciones generales: la primera que pone el acento en el *nivel o grado de ejecución* sobre algún dominio o criterio definido de antemano y que actualmente se conoce como test referidos al criterio; la segunda en la que la aproximación a los *rasgos latentes* o medidas de los atributos de las personas, interesadas en situar a las distintas personas a lo largo del continuo que define o representa el atributo; y tercera la aproximación *de procesamiento de la información* que intenta determinar los procesos cognitivos subyacentes a las respuestas dadas a los ítems (p 32-35).

### **3.3.5.7 Clasificación general de los test de Inteligencia**

Resulta complicado realizar una clasificación de los test de inteligencia, y más que ésta cumpla con los requisitos que se exigen a una clasificación, pero como

un objetivo puramente didáctico se expondrá la clasificación que presenta Cerda (1973, p. 79-80) partiendo de los métodos más importantes de exploración de la inteligencia en dos grandes grupos: a) Test de inteligencia para niños y b) test de inteligencia para adultos.

**Tabla 2. CLASIFICACIÓN GENERAL DE LOS TEST DE INTELIGENCIA PARA NIÑOS**

1° TEST TIPO BINET	BINET-SIMON 1911 BINET-TERMAN 1917 (REVISIÓN STANFORD) TERMAN-MERRIL 1937 (NUEVA REVISIÓN STANFORD)	
2° Test tipo Binet para edades preescolares (baby test)	Escalas de Gesell Escalas de Bühler Otras escalas	
3° Test de tipo Point scales	WISC (Wechsler Intelligence Scale Children)	
4° Test no verbales	Test de Goodenough Laberintos de Porteus Escala de performance de Grace Arthur	
5° Test colectivos	Verbales	Vocabularios Escala verbal de Pinter.
	No verbales	Matrices Progresivas (Raven) Test Culture Free (Cattell) Escalas No verbales (Pinter)

**Tabla 3. CLASIFICACIÓN DE TEST DE INTELIGENCIA PARA ADULTOS**

1° TESTS INDIVIDUALES	TIPO BINET TIPO POINT SCALES	TERMAN-MERRIL WECHSLER-BELLEVUE F1 WECHSLER-BELLEVUE FII WECHSLER ADULT INTELL SCALE
2°. Test colectivos	Test tipo Army Tets Test Factoriales	Test de factor g Test y baterías de factores

Finalmente para concluir este capítulo se abordará en términos generales el proceso que se sigue para traducir las conductas psicológicas, en elementos de un test, cabe aclarar que son pocos los estudios que se han hecho al respecto encontrándose algunos autores como Cronbach (1970), Roid y Haladyna (1980) quienes han tratado este tema pero centrándose casi exclusivamente en los test de rendimiento educativo (en Martínez, 1995, p.48)

En general para la construcción de un test, como instrumento de medida de un determinado constructo se basa en un modelo teórico, dentro del cual se inscribe el constructo y una serie de conductas observables ligadas a aquél mediante definiciones sistemáticas, y son éstas las que precisamente deben plasmarse en los ítems del test.

Una vez establecido lo anterior se requiere delimitar los ítems que serán incluidos en el test, ya que pretender crear todos los ítems posibles resulta ilusorio en la práctica. "Una forma de operacionalizar este procedimiento es la llamada

especificación de los ítems que incluye: fuentes del contenido del ítem. descripciones de los estímulos o situaciones problema características de las respuestas correctas y si es necesario (test de elección múltiple) de las incorrectas. También se debe atender a dos propiedades independientes de los ítems: el contenido sustantivo; y el proceso cognitivo u operaciones que el examinado debe emplear para resolver la tarea planteada en el ítem”(Martínez, 1995, p.50)

Sobre este último aspecto han surgido sistemas jerárquicos para categorizar las operaciones cognitivas que son útiles para establecer las especificaciones del test:

- 1) Conocimiento: recuerdo del material, en forma similar a como fue presentado durante las instrucciones.
- 2) Comprensión: traslado, interpretación o extrapolación de un concepto en forma algo diferente de la originalmente presentada o practicada.
- 3) Aplicación: resolver problemas nuevos por medio del uso de principios familiares o generalizaciones.
- 4) Análisis: romper un problema en sus elementos componentes utilizando un proceso que requiere conocimientos de elementos múltiples, relaciones y/o principios de organización entre estos elementos.
- 5) Síntesis: combinar elementos en un todo usando una estructura original o resolver un problema que requiere la combinación de varios principios secuencialmente en una nueva situación.
- 6) Evaluación: empleo de criterios internos (generados por el sujeto) o externos para hacer juicios críticos en términos de seguridad, consistencia, etc. (Martínez, 1995, p 50).

Otra corriente que realiza un análisis sobre el contenido de las pruebas psicométricas, lo encontramos en las posturas genéticas, las cuales consideran a su vez varios aspectos:

- La construcción de las pruebas debe considerar las variables que puedan alterar la homogeneidad de las muestras. Una variable es la edad cronológica y

ésta es fundamental en las pruebas de nivel mental. en las que se busca valorar los comportamientos inteligentes

- El análisis genético permite realizar una interpretación más rica de los procesos y las estructuras mentales, ya que se fundamenta en el método clínico desarrollado por Piaget, que busca saber cómo y por qué se piensa que cambios estructurales se ponen en juego y le permiten llegar a elaborar tal pensamiento, es decir, tratar de definir que piensa un sujeto de determinada edad en condiciones establecidas.
- Se parte del supuesto de que todo comportamiento que aparece en un nivel de desarrollo cronológico supone procesos de tipo genético, es decir, de una estructura que se construye a través de estadios (equilibración).

La psicología genética se observan realiza aportaciones que se considera conveniente mencionar sobre las pruebas psicométricas:

1. Las pruebas psicométricas llevan implícita la noción de evolución; dentro de ella encontramos dos posturas la primera que pone el énfasis en la variaciones específicas del fenotipo a partir de la presión de un medio cambiante. Dentro de esta perspectiva encontramos la definición de inteligencia de Spearman.
2. La segunda implica los factores hereditarios, que se definen por las características aparecidas por azar en los individuos seleccionados como los más "aptos" para sobrevivir. Se fundamenta desde el punto de vista psicológico en las teorías del ensayo y error. Se integra el genotipo a modo de formas a priori impresas en el neonato. Ejemplo de esta postura la encontramos en las pruebas desarrolladas por Binet

3. Integra ambas posturas define a la vida como un sistema de autorregulación progresiva que va de lo morfogenético, a lo estructural, y de ahí a lo funcional, donde las funciones cognitivas serían los órganos especializados de la autorregulación de los intercambios en el seno del comportamiento.

Esta postura, permite ver en las pruebas psicométricas las operaciones específicas que se ponen en juego a través del estímulo presentado en el ítem a resolver. Un ejemplo de ello sería -y con ello concluimos este capítulo- el papel de las pruebas verbales desde la genética será verificar si se encuentran presentes los sistemas de organización de significados en jerarquías que suponen procesos más generales de operación como son: la clasificación, la inclusión, etc (Pain, 1976, p, 12-13)

Como se puede ver, hablar sobre inteligencia y su medición es una tarea compleja y ardua debido a la cantidad de factores que intervienen para su comprensión, por lo que se espera que el capítulo haya aportado un panorama lo suficientemente amplio para ubicar el objeto central de este trabajo: la baremación del test de factor g de Cattell para México a través, del estudio empírico que se presenta en nuestro siguiente apartado

## CAPÍTULO 3

### BAREMACIÓN DEL TEST DE FACTOR "G" EN SU ESCALA TRES DE R.B. CATTELL Y A.S.K. CATTELL EN LA CIUDAD DE MÉXICO.

#### 3.1 Planteamiento de la investigación e hipótesis

El uso de los test se ha ido popularizando cada vez más en los distintos centros educación de nuestro país, principalmente en las instituciones de educación media superior y superior, a través de la creación de distintos centros de atención psicopedagógica, áreas de evaluación psicológica e incluso en los espacios de los departamentos de Orientación Educativa, donde se ha empezado a formalizar el manejo de instrumentos de medición psicológica.

El uso de diversos test para la valoración de las habilidades, aptitudes, rasgos de personalidad, intereses, etc., son utilizados para obtener diagnósticos, que permitan generar procesos de selección más objetivos, definir un problema de comportamiento o rendimiento escolar, el grado de madurez de un niño en edad escolar, entre otros, son algunas de las necesidades que ha llevado a la utilización de los test como medios para obtener mayor información sobre los sujetos de la enseñanza. Sin embargo, muchas de las pruebas utilizadas, sino es que la gran mayoría de ellas, son importadas y por lo tanto baremadas o estandarizadas en poblaciones de su lugar de origen, lo que representa un problema en cuanto a la confiabilidad que éstas puedan tener en sus resultados para la población que acude a solicitar un servicio de evaluación.

Dado que el concepto de baremación es central en este trabajo se considera conveniente definir lo que se entiende por baremación; en términos estadísticos significa acción de baremar, que implica la aplicación de un conjunto de técnicas y procedimientos que permiten construir un baremo o norma, que se define como una escala de puntuaciones obtenidas y que se usa para tipificar un test. La baremación trata de construir un marco de referencia pertinente y coherente que permita dar algún significado a las puntuaciones brutas obtenidas, con el pase o aplicación de una prueba o test. Existen tres tipos principales de normas o baremos: a) normas o puntuaciones cronológicas (edad mental, cociente intelectual); b) escalas de puntuaciones (centiles, deciles, percentiles); c) escalas de puntuaciones típicas (puntuaciones z, derivadas, etc)" (Diccionario de Ciencias de la Educación, 1985). Estos baremos o normas, nos indican la posición relativa que un sujeto ocupa respecto a la tarea realizada de acuerdo a su población de referencia.

Por otra parte se considera que la baremación difiere de la estandarización, en el sentido de que la primera solo comprende una región (por ejemplo, una ciudad o una localidad o incluso un grupo escolar) y una población (por ejemplo, escolares, adolescentes, adultos, profesionales, etcétera) específicas, siendo la muestra de estudio no completamente representativa de acuerdo a los fundamentos estadísticos, mientras que la estandarización implica un proceso más amplio, es decir abarcar una muestra representativa de la población de todo un país, mientras que la primera es más local y específica, la segunda sería más general y amplia. Es por ello que el título del trabajo se especifica como baremación ya que el test fue aplicado únicamente en la zona metropolitana de la ciudad de México, en estudiantes de educación Media Superior y Superior de escuelas públicas.

En México son pocos los trabajos que se han realizado en cuanto a la baremación y/o estandarización de test, actividad que se encuentra descuidada a pesar de ser una de las fuentes de trabajo más importantes de un gran número de psicólogos e instituciones de nuestro país. Conocedores de esta situación la Facultad de Estudios Superiores Iztacala, funda en 1994 la Unidad de Evaluación Psicológica (UEPI), con la intención de brindar un servicio de evaluación psicológica a la sociedad en general, incluyendo dentro de este proyecto tres ejes fundamentales que orientan el trabajo de la Unidad a saber: Servicio, Investigación y Docencia; cada uno de ellos con sus objetivos específicos, siendo en el segundo de éstos donde se inserta el trabajo desarrollado en esta tesis.

Los objetivos específicos que se plantearon en el proyecto general sobre el rubro de investigación fueron los siguientes:

- 1) Evaluación de la calidad del servicio. Tiene como finalidad detectar los niveles reales de la calidad del servicio de evaluación que se ofrece.
- 2) Diseño y construcción de instrumentos de evaluación. Esta línea de investigación se enfoca a la identificación de los instrumentos de evaluación con los que cuenta el servicio, así como la construcción de aquellos que no existen en el mercado.
- 3) Confiabilización de instrumentos. El objetivo de esta línea es realizar estudios donde se pueda investigar el grado de exactitud con los que se miden los instrumentos utilizados en la UEPI.
- 4) Baremación y/o estandarización de instrumentos. Su finalidad es obtener los baremos propios de las pruebas más utilizadas en la Unidad. En la

estandarización se exploran las condiciones más adecuadas para usar un determinado instrumento así como las normas más pertinentes para comparar las puntuaciones obtenidas en un determinado individuo.

- 5) Validación de instrumentos. Se realizarán investigaciones dirigidas a determinar la validez de instrumentos, evaluar el atributo el concepto o habilidad.

Buscando ampliar cada uno de los objetivos planteados en el proyecto de la UEPI, los estudiantes tanto de licenciatura como de posgrado que se inscriben al proyecto, a parte de involucrarse en el programa de atención a usuarios, se define de acuerdo a sus intereses, la exploración de posibles líneas de investigación que se desprendan de la misma práctica que se va realizando. En el caso particular de la presente investigación la atención se centró en una de las pruebas de inteligencia más utilizadas en la UEPI, el test de factor "g" de Cattell (1994) en su escala 3, por ser una prueba que presenta ventajas en cuanto a su aplicación; de corto tiempo y de tipo colectiva y porque se aplica a un promedio de 150 estudiantes por año, lo que significa un índice alto para la UEPI en cuanto a su utilización, ya que proporciona información muy útil sobre tres capacidades cualitativas de cognición en grandes grupos de población, además de reducir en gran medida la influencia de otros factores tales como fluidez verbal, nivel cultural, clima educacional, etc. (Cattell y Cattell. 1994)

En la actualidad y desde siempre, la inteligencia es uno de los constructos que conforman un gran porcentaje de peticiones de evaluación psicológica, razón por la cual se decidió llevar a cabo este estudio, con la intención de obtener nuestros propios baremos, ya que como se mencionó con anterioridad la prueba

fue baremada en España, con población estudiantil 8º de E.G.B, 1º, 2º y 3º de BUP y 1er Año de COU, de una región específica del país. Aunque los resultados obtenidos en tal baremación son válidos y confiables para este tipo de población, se requiere la baremación específica para poder evaluar inteligencia de una manera más confiable en nuestra población, comparando a los jóvenes evaluados con otros jóvenes de su misma idiosincrasia, nivel educativo y cultural, ya que las interpretaciones que obtenemos con las tablas construidas en España no concuerdan con el desempeño académico de nuestros estudiantes y los ubica por debajo de su media, por lo tanto el objetivo general de esta investigación se centrará en lo siguiente:

### **3.2 Objetivo General**

Obtener los baremos de la prueba de inteligencia Factor "g" en su escala 3 de R.B. Cattell, con el fin de que los resultados obtenidos de la muestra elegida, sean válidos para sus posteriores aplicaciones en poblaciones de las mismas características de la zona metropolitana de la ciudad de México.

La investigación es de tipo descriptivo<sup>3</sup> ya que específicamente buscó obtener los baremos correspondientes y que a su vez, se cumpliera uno de los objetivos del proyecto de evaluación que es lograr en la medida de lo posible la adecuación de dichos instrumentos y que éstos sean realmente útiles para el servicio que se brinda a la comunidad.

---

<sup>3</sup> El objetivo central de los estudios descriptivos, es obtener un panorama más preciso de la magnitud del problema o situación, jerarquizar los problemas, derivar elementos de juicio para estructurar políticas o

### **3.3 Hipótesis de trabajo**

Se espera que los resultados obtenidos en la baremación del Cattell en población mexicana muestren diferencias en los niveles de ejecución, en comparación a los baremos de la población española, a pesar de que este test mide inteligencia general libre de cultura.

### **3.4 Método**

#### *3.4.1 Sujetos*

En este estudio la muestra fue seleccionada de manera aleatoria, siguiendo los mismos criterios establecidos en el manual de la prueba publicado por TEA (1994), con grupos equivalentes en lo que se refiere al nivel educativo. La muestra estuvo conformada por 1322 alumnos de 34 escuelas públicas de Educación Media Superior (Bachillerato) y Educación Superior, de la zona metropolitana de la Ciudad de México. En la siguiente tabla se muestra como quedó constituida la misma:

---

estrategias operativas, conocer las variables que se asocian y señalar los lineamientos para la prueba de la hipótesis (Rojas 1982 p 31)

Tabla 4 Conformación de la muestra de estudio

N = 1322				
H = 592 44.78%		M = 708 53.56%		NR = 21 1.59%
34 Escuelas Públicas				
Grado Escolar	22 Educación Media Superior		12 Educación Superior	
	10 Vespertino	12 Matutino	6 Vespertino	6 Matutino
1°.	3	4	6	6
2°.	3	4		
3°.	4	4		
	N = 895 67.70%		N = 427 32.30%	

Como se aprecia en la Tabla 4, del total de la muestra, 592 fueron hombres (44.78%), 708 mujeres (53.56%) y 21 no reportaron el dato (NR=1.59%). Del total de la muestra (1322 alumnos), 895 fueron elegidos de los tres años del ciclo de Bachillerato (67.70%) y 427 del primer año de estudios universitarios (32.30%).

La elección se realizó de la siguiente manera. Del listado de escuelas públicas de educación media superior y educación superior, se escogieron aleatoriamente 22 y 12, considerando en número de instituciones existentes en las diferentes delegaciones políticas y municipios conurbados que conforman la zona metropolitana de la ciudad de México, eligiendo al menos una escuela de cada delegación y municipio respectivamente. Debido a que son más las instituciones de escuelas de educación media superior existentes se consideró necesario abarcar al menos una de cada sistema (Colegios de Bachilleres, CCH, CECYT,

CONALEP etc). y para educación superior al menos dos carreras que se ofertan en los sistemas públicos (UNAM, IPN, UAM, Institutos Tecnológicos Regionales, UAEM). Para llevar a cabo el estudio; de las primeras escuelas se seleccionaron también aleatoriamente, 10 del turno vespertino y 12 del matutino y de las segunda, seis y seis, respectivamente, considerando aquellas instituciones que manejen ambos turnos, con la finalidad de abarcar en la muestra al mayor número de sujetos que se encuentran estudiando en estos niveles. Para tener una mejor visión de la composición de la muestra a continuación se presenta en las tablas 5 y 6 el total de instituciones seleccionadas para el estudio:

Tabla 5. Instituciones de educación media superior seleccionadas para el estudio.

No.	Nombre de la escuela	Semestre	Turno
1	Escuela Nacional Preparatoria No. 4	Quinto	Matutino Vespertino
2	Escuela Nacional Preparatoria No. 5	Quinto	Matutino vespertino
3	Escuela Nacional Preparatoria No. 9 "Pedro Alva"	Tercero Quinto	Vespertino Matutino
4	CCH Naucalpan	Segundo y cuarto	Matutino
5	CCH Sur	Segundo y cuarto	Vespertino
6	CCH Oriente	Segundo, cuarto y sexto	Vespertino
7	Colegio de Bachilleres No.12 "Netzahualcoyotl"	Sexto	Vespertino
8	Colegio de Bachilleres No. 1 "El Rosario"	Primero  Segundo	Matutino Vespertino
9	Colegio de Bachilleres No. 2 "100 metros"	Primero	Matutino
10	CECyT No. 2 "Miguel Bernal"	Segundo Sexto	Vespertino
11	CECyT No. 8 "Narciso Bassols"	Cuarto Segundo	Matutino Vespertino
12	CECyT No.15 "Diodoro Antunas Echegaray"	Cuarto y segundo	Vespertino
13	Preparatoria Oficial Atizapan No. 64	Segundo	Vespertino
14	Preparatoria Oficial Tultitlán No.11	Sexto	Vespertino
15	CETIS No. 5	Segundo, Cuarto sexto	Matutino Vespertino
16	CETIS No. 95	Segundo	Matutino
17	CBTIS No. 160	Sexto	Vespertino
18	CBTIS No. 202	cuarto y sexto	Matutino
19	CONALEP GUSTAVO A. MADERO	primero segundo y tercero	Matutino
20	CONALEP ATIZAPAN	primero y cuarto	Matutino
21	CONALEP NEZA III	Tercero	Vespertino
22	CONALEP MEXICO-CANADA	Quinto	Matutino

Tabla 6. Escuelas de Educación Superior

No.	Nombre de la escuela	Carrera	Turno	Semestre
1	Tecnológico Regional de Tlalnepantla	Administración	Matutino	Primero
2	Escuela Superior de Turismo (IPN)	Turismo	Matutino	Segundo
3	ESUQUIE-IPN	Ing. Química	Matutino	Primero
4	UAM-Azcapotzalco	Arquitectura D. Industrial y D. Gráfico	Matutino	Primero
5	UAM-Iztapalapa	Carreras de Ciencias Biológicas	Matutino	Primero
6	UAM-Xochimilco	Carreras de Ciencias Sociales	Vespertino	Primero
7	FES-Cuautitlán	Contaduría	Vespertino	Segundo
8	ENEP-Aragón	Ing. Mecánicas y Eléctricas	Matutino	Segundo
9	ENEP-Acatlán	Pedagogía y Actuaría	Vespertino	Segundo
10	FES-IZTACALA	Optometría	Matutino	Segundo
11	UAEM	Economía y Relaciones Internacionales	Matutino	Segundo
12	FES-Zaragoza	Ing. Química	Matutino	Segundo

Finalmente, de cada escuela de educación media superior, se seleccionó un grupo de alguno de los tres grados escolares que conforman el ciclo de bachillerato y, de las escuelas de educación superior, se seleccionó también un grupo de tal forma que cada uno de ellos fuera de diferente carrera. Cabe aclarar que esos grupos de alumnos seleccionados no fueron de manera aleatoria, sino que, dadas las características elegidas (grado escolar o carrera), el grupo fue asignado por el encargado del plantel de acuerdo a que el profesor cediera su hora de clase para poder aplicar la evaluación. Es por esta razón que la muestra no estuvo balanceada con respecto al sexo.

De esta manera, la aplicación de la prueba de Cattell se llevó a cabo, en cada grupo de alumnos seleccionados, dentro de su plantel, salón y pupitre correspondientes a la hora de la aplicación. Se les dieron las instrucciones de aplicación como están impresas en el Manual y una vez terminado el tiempo de aplicación se procedió a recoger y calificar la prueba.

### 3.3.2 Instrumento de medida

#### *Descripción del test de factor "g" en su escala 3*

En este test Cattell y Cattell (1994) están interesados en medir la inteligencia; a la que conciben como la capacidad mental general o "factor "g" Considera que subyace a las aptitudes mentales primarias y prácticamente a casi todas las conducta o tareas específicas. La pretensión de Cattell es responder a la necesidad de evaluar con precisión la inteligencia de sujetos con diferente lenguaje, cultura y nivel social o educativo, reduciendo al máximo los efectos debidos a estas diferencias. Aunque para ellos la inteligencia es muy importante, consideran que solamente es un elemento de una larga lista de cualidades individuales que es necesario considerar cuando se intenta comprender y predecir ampliamente el comportamiento humano.

Las investigaciones sobre educación han demostrado frecuentemente que la inteligencia es un importante factor para predecir los resultados académicos, pero también se ha descubierto que con la inclusión de algunos factores de personalidad y motivación en las baterías de evaluación psicológica, el nivel de exactitud de la predicción mejora notablemente.

Además de lo anterior, se escogió esta prueba debido a que constituye un instrumento de gran utilidad para evaluar la inteligencia, ya que puede ser aplicado a una serie de situaciones en las que resulta muy recomendable su utilización. La Escala 3 está diseñada para ser aplicada a estudiantes que se encuentren en los

últimos cursos de enseñanza media superior y universitarios, así como población general de adultos con un nivel relativamente elevado de cultura, esto es, con formación mínima de secundaria; su aplicación es colectiva, durando aproximadamente media hora y es fácil de calificar a través de una plantilla, y también debido a que los autores mencionan que algunas de sus finalidades son: determinar el potencial de un candidato para realizar tareas en las que esté implicada la aptitud cognitiva, seleccionar dentro de una clase a los estudiantes más capacitados para participar en programas intensivos de instrucción, orientar a los estudiantes respecto a su posible éxito en el colegio o a la obtención de becas y aumentar la efectividad de la orientación vocacional.

La Escala 3 de inteligencia de Cattell está diseñada para apreciar la inteligencia individual por medio de pruebas no verbales, de tal manera que para su realización se requiere solamente que el sujeto perciba la posibilidad de relación entre formas y figuras. Esta escala está formada por cuatro subtests: series, clasificación, matrices y condiciones, que implican contenidos perceptivos distintos con objeto de evitar que determinadas diferencias de esta aptitud perceptiva influyan en los resultados de la medida de inteligencia.

De acuerdo con Martínez (1995, p. 32), esta prueba se clasificaría dentro de los test psicométricos que presentan las siguientes características: "Evalúan las respuestas según normas cuantitativas; Todos sus elementos se valoran de forma numérica e independientemente; el resultado final es una puntuación cuantitativa; Suelen referirse a características psicológicas concebidas como unidades más o menos independientes, es decir, suelen medir rasgos diferenciados del sujeto".

Por su finalidad es considerado tanto como un test de investigación como de diagnóstico; por el tipo de planteamiento de los problemas que maneja se considera que es una prueba de ejecución máxima, donde la situación de cada elemento plantea un problema que el sujeto ha de resolver poniendo en funcionamiento su capacidad en alto grado. Por su modalidad de aplicación es de los llamados de papel y lápiz, ya que admite un formato de presentación, protocolos, cuadernillos y hojas de respuesta en forma escrita. Es un test que puede ser aplicado tanto en forma individual como colectiva. Demanda del sujeto rapidez y potencia en sus respuestas, ya que el tiempo de ejecución está rigurosamente controlado esperando valorar la rapidez en la respuesta del sujeto, a la vez que se requiere de calidad en las ejecuciones.

### **3.5 Análisis Estadístico**

Para el análisis estadístico de los resultados obtenidos en la aplicación de la prueba de inteligencia de factor "g" de R.B. Cattell y A.K. Cattell, se consideraron los mismos procedimientos estadísticos utilizados en la baremación hecha en España, con la finalidad de obtener una indicación global de la confiabilidad de la prueba en su ámbito de aplicación, es decir, si esta prueba está siendo utilizada en la UEPi como una prueba de orientación que da al evaluador elementos para determinar el nivel de inteligencia del usuario que atiende, la información que se obtenga debe ser más precisa y por consiguiente más útil para la interpretación del caso particular relacionándolo siempre con el grupo de las mismas características, de aquí que sólo se pretenda obtener nuestras propias tablas de interpretación. A

continuación se describe de manera general los procedimientos estadísticos que se siguieron en la baremación de la prueba en España.

Una de las características fundamentales de una prueba es la confiabilidad, estadístico que indica la precisión (consistencia o estabilidad) de los resultados y señala el grado en que las medidas están libres de errores aleatorios. Fernández y Carroles (en Silva, 1993, p. 75) afirman que "los teóricos de la medida denominan confiabilidad a la consistencia con que mide un instrumento ( ... ) la seguridad es el grado de acuerdo o correspondencia entre medidas, que deberían ser iguales sino existiesen errores de medición. Mientras Kerlinger (en Silva, 1993) considera que la confiabilidad de un instrumento dependerá del grado en que produzca errores en la medición, dicho de otro modo, la confiabilidad se puede definir como la carencia relativa de errores de medición del instrumento. La confiabilidad es la exactitud o precisión de un instrumento de medición.

Normalmente la confiabilidad se expresa en forma de coeficiente de correlación. Para interpretar los coeficientes de correlación se debe comenzar por restringir su significación, es decir, que estas deben ser significativamente diferentes de cero, ya que si esta diferencia no surge, justamente no hay nada que interpretar. Pero si también difieren significativamente de cero, la interpretación dependerá del propósito para que han sido calculadas. (Smith, 1971) En el caso de la baremación de esta prueba se pretende predecir la posición probable en una de dos distribuciones correlacionadas (calificaciones obtenidas en el curso anterior a la aplicación de la prueba) a partir de la posición de otra, que en este caso es el test de inteligencia mencionado.

La confiabilidad de la prueba se estima comúnmente correlacionando consigo misma mediante de una de las tres maneras siguientes: una prueba con la prueba efectuada otra vez una forma con una forma equivalente, o una mitad con la otra mitad. A los valores derivados de estas correlaciones se les llama confiabilidad y se designa con el símbolo  $r$ . En el caso específico de la prueba en cuestión se utiliza el tercer procedimiento mencionado, al que también se le conoce como el método de "dos mitades" o aproximación "non-par" que se obtiene usualmente correlacionando renglones nones con renglones pares, en lo que se refiere a la prueba se correlacionan el total de aciertos con los elementos pares con el total de aciertos en los elementos impares. Los resultados obtenidos se corregirán con la fórmula de Spearman-Brow, siendo la siguiente:

$$R_{11} = \frac{2r_{1/2}}{1 + r_{1/2}}$$

Donde  $r_{11}$  es el coeficiente de confiabilidad estimado para la prueba completa y  $r_{1/2}$  es el valor obtenido correlacionando las dos mitades.

Otro aspecto a considerar es su validez, ya que esta indica el grado en que la prueba cumple con los cometidos para los que fue diseñada. Esto implica cierta dificultad debido a la búsqueda de criterios adecuados que indiquen lo que la prueba está tratando de medir, fuera de la prueba misma, con la cual la prueba pueda ser comprobada por correlación. En el caso de la prueba objeto de este estudio, algunos criterios que se tomaron en cuenta, fue el aprovechamiento escolar, el proceso de maduración intelectual criterio que ha sido corroborado repetidas veces con la aplicación de test de inteligencia en sujetos de diversos grupos de edad. Otro aspecto que también ha reportado diferencias sobre observaciones realizadas en los resultados de la baremación de este tipo de

pruebas es el relacionado con las diferencias intersexos. Las observaciones que se han hecho de muestras amplias y muy homogéneas, como la que se ha tomado para la investigación, le da un mayor grado de validez al instrumento.

El hecho de que ya haya sido validada es otro elemento que permite unificar criterios respecto a su validez. "Es costumbre considerar válida una prueba si se encuentra una alta correlación entre la prueba y un criterio independientemente satisfactorio, con tal de que las confiabilidades como de la prueba como del criterio sean satisfactorias" (Smith, 1971, p. 156). Aunque no necesariamente se espera que la correlación entre una prueba y su criterio sea tan grande como la esperada en una autocorrelación, o coeficiente de confiabilidad.

### **3.5.1 Procedimiento seguido para la Baremación.**

1. Una vez obtenida la aplicación total de la prueba a la muestra, se generó una base de datos en SPSS versión 6.01, para su tratamiento estadístico, se eligió este programa debido a que a través de él se pudo "hacer más rápido y eficiente el análisis cuantitativo de los datos" –además de considerar que –"es el programa más ampliamente conocido y usado debido a que proporciona un poderoso sistema de análisis estadístico y de administración" (Silva, 1998, p. 17).
2. Para obtener las puntuaciones centiles<sup>4</sup> se aplicaron los estadísticos básicos, obtenidos automáticamente por el programa, para la elaboración de los centiles se siguió el siguiente procedimiento (descrito en Morales, 2000):

---

<sup>4</sup> Los rangos percentiles o también llamados centiles, constituyen una de las calificaciones más utilizadas en las tablas de normas de pruebas estandarizadas, porque permiten en general una mejor interpretación de los resultados, aún cuando tienen ciertas limitaciones. El percentil es cualquiera de los 99 puntos de igual amplitud.

- a) Se ordenaron todas las calificaciones crudas en orden de menor a mayor.
  - b) Se encontró la frecuencia de ocurrencia de cada calificación
  - c) Se halló la frecuencia acumulada a través de cada calificación
  - d) Se encontró el punto medio de la frecuencia acumulada de cada calificación.
  - e) Se convirtieron a porcentajes acumulados el punto medio de la frecuencia acumulada.
  - f) Se sustituyó el valor de la calificación estándar normalizada correspondiente por su porcentaje acumulado en una distribución normal de probabilidad
  - g) Finalmente de los porcentajes acumulados se encontraron los rangos percentiles , completando los valores de los puntos medios a números enteros a excepción del 1 para el 0 y 99 + para el 100
3. Una vez obtenidas las puntuaciones centiles, se procedió a obtener los cocientes intelectuales, para lo cual se obtuvieron las puntuaciones Z, debido a que las puntuaciones tipificadas z se expresan en decimales y, aproximadamente la mitad de ellas son negativas, es necesario hacer una transformación lineal que consiste en aumentar la escala de modo que desaparezcan las puntuaciones negativas y emplear una desviación típica mayor, de forma que se eviten los decimales (Downie y Heat. 1973).

Aplicando según lo establece el manual para obtener nuestras propias tablas, si se cumple la condición de disponer de un número de casos suficientemente grande, (condición consideramos se cumple), se iguala la media y la desviación típica de las puntuaciones directas que obtenga a 100 y 16 respectivamente (Cattell y Cattell, 1994).

---

en que se divide una distribución de frecuencia dentro de un grupo de 100 El rango percentil es la posición

### 3.6 Resultados

Una vez terminado el tratamiento estadístico de los datos se obtuvo la tabla de baremos, que incluye tanto los percentiles, como los cocientes intelectuales correspondientes para cada grupo estudiado. A nosotros nos interesaba obtener en particular los baremos de esta prueba específicamente en población de bachillerato, puesto que ellos son los que más tarde cursarán una carrera universitaria, y en primer año de universidad, por las evaluaciones que nos piden para el Programa de Alta Exigencia Académica de la Universidad Nacional Autónoma de México, con la finalidad de poder predecir rendimiento escolar; por eso es que tanto la muestra elegida como los baremos obtenidos son para esas dos poblaciones, de bachillerato -1o., 2o y 3o años-, que correspondería a una edad de 15, 16 y 17 años, respectivamente y de primer año de universidad, con una edad aproximada de 18 años.

Por otro lado, en las tablas de la baremación española, para las puntuaciones centiles se tienen cuadros para Formación profesional, 8o. de EGB y 1o. y 2o. de BUP, 3o. de BUP, COU y población general de adultos y, para las puntuaciones de CI, cuadros para 13, 14, 15 y 16 y más año

---

relativa de un individuo dentro de un grupo específico (Morales 2000, p 58)

Tabla 6 Submuestras seleccionadas de las poblaciones española y mexicana

TABLAS DE CENTILES				
Baremos para la población española				
Formación profesional	8o. de EGB, 1o y 2o. BUP	3o. BUP	COU	Población general adulta
Baremos para la población mexicana				
---	1o y 2o. y 3o Bachillerato	1o. Universidad	---	---
TABLAS PARA EL COCIENTE INTELECTUAL				
Baremos para la población española				
13 años	14 años	15 años	16 y más años	
Baremos para la población mexicana				
---	---	1º., 2o. Y 3o Bachillerato	1o Universidad	

Como se observa en la Tabla 6, en ambas poblaciones las submuestras son diferentes, por lo que no pueden existir comparaciones equivalentes. De hecho, esta era una limitante para los alumnos a evaluar del primer año de universidad. aparte de ser las puntuaciones de una población diferente a la nuestra Sin embargo, tomando como equivalentes para las puntuaciones centiles al Bachillerato de México con los puntajes del 8o de EGB y 1o y 2o. de BUP en España, así como al primer año de universidad en México con el 3o. de BUP español, tenemos:

Tabla 7. Puntuaciones brutas y centiles para las poblaciones de España y México

<i>Puntuaciones centiles</i>	PUNTUACIONES BRUTAS			
	8o de EGB y 1o y 2o BUP España	1o, 2o, y 3o Bachillerato México	3o BUP España	1o Universidad México
99	30 - 50	30 - 50	32 - 50	31 - 50
75	22	23	24	23
50	20	20	22	21
30	17	18	20	18
1	0 - 11	0 - 9	0 - 13	0 - 9
	N = 1298	N = 895	N = 478	N = 427
Media	19.77	20.8	22.09	21.3
Desv Estándar	4.25	4.75	4.00	4.59

Como podemos observar al comparar ambas poblaciones, no existen diferencias muy grandes entre las puntuaciones brutas requeridas por una y otra población para alcanzar determinados centiles; apenas son del orden de dos aciertos cuando más, incluyendo las medias de las poblaciones. Sin embargo, hay que recordar que las submuestras no son equivalentes en cuanto a nivel de estudios. Por otro lado, si comparamos las puntuaciones brutas requeridas para alcanzar determinados puntajes de CI, tenemos lo siguiente:

Tabla 8. Puntuaciones brutas y Coeficiente Intelectual para las poblaciones de España y México

<i>Coefficiente Intelectual</i>	PUNTUACIONES BRUIAS			
	15 años España	1o, 2º y 3º Bachillerato México (15, 16 y 17 a.)	16 años y más España	1o. Universidad México (18 y más a.)
148	33	35	35	35
132	29	29	31	30
116	26	26	27	26
<b>100</b>	<b>22</b>	<b>22</b>	<b>23</b>	<b>21</b>
84	17	16	19	16
68	13	12	15	12
52	9	9	11	9

Como observamos aquí, al igual que en las puntuaciones mostradas en la tabla anterior, tampoco son muy grandes las diferencias, en las puntuaciones de las poblaciones estudiadas, sobre todo en la primera comparación. Sin embargo, en la segunda medición de la tabla 8, excepto en el CI de 148, en todos los demás CI fueron sistemáticamente mayores a los puntajes de la población española.

Finalmente, se presenta en la tabla 9, los baremos, en donde se muestran las puntuaciones generales para estudiantes mexicanos, varones y mujeres de la zona metropolitana de la Ciudad de México.

Tabla 9 BAREMOS DEL TEST CATTELL PARA ESTUDIANTES MEXICANOS (VARONES Y MUJERES)  
FACTOR "g" 3-A

Educación Media Superior			Educación Superior	
Centil	Puntuación Bruta	Cociente intelectual (IQ)	Puntuación Bruta	Cociente Intelectual (IQ)
99	30-34	130-145	31-35	134-148
98	29	128	30	130
97	28	124	29	128
96	-	-	-	-
94	-	-	28	124
93	27	121	-	-
92	-	-	27	121
89	26	117	-	-
88	-	-	26	117
84	25	113	25	113
77	24	110	-	-
76	-	-	24	110
71	23	107	-	-
69	-	-	23	107
64	22	103	-	-
60	-	-	22	103
54	21	99	-	-
50	-	-	21	99
45	20	96	-	-
42	-	-	20	96
37	19	93	-	-
34	-	-	19	93
30	18	90	-	-
24	-	-	18	90
23	17	86	-	-
16	16	83	17	86
12	-	-	16	83
11	15	79	-	-
8	14	75	15	79
6	13	72	14	75
5	-	-	13	72
4	12	69	-	-
3	11	66	12	69
2	10	62	10-11	62-66
1	0-9	58-28	6-9	48-58
N	895		4.27	
Media	20.8		21.3	
Desviación típica	4.79		4.59	

## CONCLUSIONES

Al término de esta investigación, cuya finalidad fue obtener la baremación del test de inteligencia factor "g" de Cattell, en población mexicana, los resultados obtenidos nos muestran, en general, que las diferencias presentadas en lo que respecta a las puntuaciones medias de los grupos estudiados son apenas de uno o dos puntos. Sin embargo, en la dispersión de las puntuaciones extremas por encima o por debajo de la media si se observan diferencias significativas sobre todo en lo que respecta a las puntuaciones tope obtenidas por cada grupo. En el caso de los españoles, se tuvo estudiantes que obtuvieron hasta cincuenta aciertos (que representa la puntuación máxima en la ejecución de la prueba), mientras que en la muestra mexicana la mayor puntuación llegó a treinta y cinco aciertos.

Estos resultados confirman nuestra hipótesis que preveía que los estudiantes mexicanos obtendrían puntuaciones por debajo de la muestra española. En otras palabras, si se interpretan literalmente estos resultados se podría decir que ¿son más inteligentes los estudiantes españoles que los estudiantes mexicanos?. Esta pregunta amerita realizar una serie de consideraciones: por ejemplo; ¿tiene alguna influencia el tiempo o velocidad de reacción de un sujeto para realizar un test con control de tiempo?, o ¿un test libre de cultura supone que todos los sujetos responderán eficazmente en el tiempo predeterminado?.

Al respecto es conveniente señalar la opinión de Anderson (2001) quien afirma que el problema de desempeño de un sujeto en un test de inteligencia se centra en el tiempo de reacción (TR) que un sujeto presenta a la hora de realizar la tarea donde la velocidad de procesamiento es la fuente primordial del factor g "Por lo tanto, la velocidad a la que una persona puede procesar información más sencilla debe poder predecir su exactitud en tareas de razonamiento complejo" (p.194) Esto me lleva a la reflexión de que posiblemente el problema que presentan nuestros estudiantes a la hora de realizar pruebas en las que se mide el tiempo de reacción es la falta de entrenamiento para manejarse en actividades en las que se requiere dar solución a los problemas de manera rápida y acertada. Sobre este último aspecto el autor menciona el tiempo de inspección (TI) que explica la habilidad del sujeto para resolver el problema, y otro fenómeno que puede influir en la velocidad y exactitud para resolver tareas de razonamiento complejo es el de los potenciales evocados (PE) que tienen relación directa con la conductividad neuronal. Si regresamos a los resultados de la baremación del Cattell en cuanto a las puntuaciones tope obtenidas en la población mexicana, pueden ser útiles para explicar dichas diferencias los tiempos de reacción y los tiempos de inspección de Anderson pues estudios transculturales han demostrado que existen diferencias en la importancia del tiempo y la rapidez con que diferentes grupos culturales ejecutan tan tareas simples como cobrar una mercancía, o atender a un cliente hasta tareas complejas como resolver un problema matemático poco estructurado, comprender instrucciones complejas o generar diferentes alternativas de solución a un problema

Hay que recordar aquí lo que ya se sabe en torno a la discusión de las pruebas generales de inteligencia: que aunque los diseñadores de estas pruebas

sostengan que están libre de factores culturales y educativos, los hechos han demostrado que siempre existirán factores externos, aún además de que las prueba psicométricas son limitadas y por lo tanto no son perfectas. Se podría decir entonces, respondiendo a las preguntas formulada más arriba, que es posible que nuestros estudiantes sean más lentos que los españoles pero no por eso, menos inteligentes.

Sin duda es claro que estas reflexiones puedan generar polémica y duda, pero ¿acaso toda la historia de la inteligencia humana y su medición no ha sido una constante polémica por demás fascinante?. Aquí sólo pretendo dar una modesta opinión sobre el tema.

En cuanto a la prueba de inteligencia de Cattell, podemos decir que, a pesar de las observaciones arriba apuntadas, aún ofrece una medida confiable de las capacidades intelectuales del sujeto. Sin embargo, por su continuo uso en las *instituciones educativas de nuestro país, resultaba ya urgente contar con los baremos adecuados a nuestra población. Incluso, el hecho de que la media de ambos grupos esté cercana a la norma española (nivel bachillerato: España 20 y México 20; primer año de estudios universitarios: España 22 y México 21),* significa que aún el haber evaluado a los usuarios de la Unidad de Evaluación Psicológica, con los baremos españoles, no se afecto en lo que respecta, al menos en las puntuaciones medias. Por lo tanto, aunque es necesario investigar el por qué de ciertos resultados en la baremación con población mexicana es innegable que es una ventaja el contar con los propios.

En México, la utilización de baremos con poblaciones de otros países con culturas diferentes es común, ya que en nuestro país es casi inexistente la realización de investigaciones para la construcción tanto de nuevas pruebas como de la baremación de las existentes. El asunto no tendría mayor relevancia sino fuera porque estos tests son continuamente utilizados para tomar decisiones y alternativas de acción, como son; la contratación de personal en una empresa, la selección de estudiantes a una carrera o para otorgar una beca. Por ello es importante que se actúe con responsabilidad ya que el valor que se le asigna a los resultados obtenidos a través del uso de test, son asumidos casi literalmente. Por lo tanto, el compromiso para "quienes se desempeñan en la psicometría es contribuir a que cada vez estos instrumentos sean de mejor calidad y cumplan con las características psicométricas que garantizan su confiabilidad y validez y esto solo se logra cuando se construye y espacio académico para su desarrollo, reflexión y debate" (Rodríguez, 2001)

En el futuro, la inteligencia seguirá siendo objeto de estudio y medición, pero seguramente, lo será desde una perspectiva multidisciplinaria. En esta misma línea de pensamiento y ante los cambios profundos de las sociedades actuales, Gardner (2001) señala que " nos encontramos ante una clara disyuntiva: o bien continuamos con las nociones tradicionales de la inteligencia y de cómo se debe medir, o bien desarrollamos una manera distinta y mejor de conceptual el intelecto humano" (p.15). Uno de estos cambios que ya se están dando en la participación de investigadores de disciplinas muy distantes como la ingeniería genética, las neurociencias y la antropología, las cuales desde hace un poco más de dos década de asombrosas explicaciones

sobre el funcionamiento inteligente, han dejado en claro que la inteligencia ya no es más patrimonio exclusivo de la psicología, lo que sin duda dará como resultado una concepción más versátil, diversificada y multireferenciada del origen y naturaleza de la inteligencia humana.



## REFERENCIAS

- Adkins, D. (1996). *Elaboración de test*. México: Trilla.
- Anastasi, A. (1980). *Tests Psicológicos*. Madrid: Aguilar.
- Anderson, M. (2001). *Desarrollo de la Inteligencia*. México: Oxford
- Baltes, P. (1992). Notas sobre el concepto de inteligencia, en R. J. Sternberg y D. K. Detterman (Coord.) *¿Qué es la Inteligencia? Enfoque actual de su naturaleza y definición*. Madrid: Pirámide.
- Barbero. (1994). *Psicometría. Definiciones*.  
<http://www.umh/asignatura.unal.edu.co/psicología/programas/4563301.htm>.
- Blauberg. (1986) *Diccionario de filosofía*. México: ediciones Quinto Sol.
- Calvin H. W. (1994) Aparición de la Inteligencia. *Investigación y Ciencia*. Diciembre, 79-85.
- Carroll, B. J. (1988). ¿Qué es la inteligencia? En R. J. Sternberg y D. K. Detterman, (Coord.) *¿Qué es la inteligencia? Enfoque actual de su naturaleza y definición*. España: Pirámide
- Carroll, B. J. (1994). La medición de la inteligencia. En R. Sternberg (Ed.). *Inteligencia humana, I. La naturaleza de la inteligencia y su medición*. Barcelona: Paidós.
- Castelló, A. (1995). Modelos de inteligencia y modelos de la instrucción, realidades teóricas y funcionales. En C. Genovard, J. Bellán y F. Rivas. *Psicología de la instrucción III, nuevas perspectivas*. Madrid: Síntesis Psicológica.
- Cattell, R. B. (1989). *Test de factor "g" - Escala 1*. Manual. Madrid: TEA
- Cattell, R. B. y Cattell, A. K. S. (1994) *Tests de Factor "g", Escalas 2 y 3*. Manual. Madrid: TEA.
- Cerda, E. (1973). *Psicología Aplicada*. España: Heider.
- Díaz, J. E. (1996) De la teoría de los test a los test adaptativos computarizados: Una revisión. México. *Revista de psicología contemporánea*, año 3. Vol 2. Septiembre. P.54-61
- Diccionario de Ciencias de la Educación (1985). México: Diagonal Santillana.
- Downie, N.M. y Heath, R.W. (1973) *Métodos estadísticos aplicados*. México: Haría.
- Gardner, H. (1995). *Estructuras de la mente. La teoría de las inteligencias múltiples*. México: FCE
- Gardner, H. (2001). *La inteligencia reformulada. Las inteligencias múltiples en el siglo XXI*. Barcelona: Paidós
- García, E. (1993). *Introducción a la psicometría*. España: SXXI
- González, L. (2000) *Teorías científicas y teorías implícitas de la inteligencia: sus repercusiones en la escuela*. México. ITESM Tesis Maestría en Educación.
- Jensen, R. A. (1988) Inteligencia: definición, medida y futura investigación. En R. J. Sternberg y D. K. Detterman, (Coord.) *¿Qué es la inteligencia? Enfoque actual de su naturaleza y definición*. España: Pirámide.
- Luján, J. L. (1991). La psicofísica y el origen de los tests de inteligencia. *Revista de Historia de la Psicología*. Vol 12. no. 1, pp. 41-55.

- Martínez A. R. (1995). *Psicometría Teorías de los tests Psicológicos y Educativos*. España: Síntesis Psicológica.
- Melía, J. L., J. M. Tomás y A. Oliver (1992) Perspectivas históricas de la teoría de la Respuesta al Ítem. *Revista de historia de la psicología*. Vol. 13 No 1, pp.105-129
- Morales, L. (2000). *Psicometría Aplicada*. México: Trillas.
- Mora, J.A. (1993). Estudiemos la inteligencia como proceso básico *Psicología General y Aplicada* 46 (3) pp. 285-291
- Muñiz, J. (1990). *Teoría de la respuesta a los ítems Un nuevo enfoque en la evolución psicológica y educación*. España: Pirámide
- Muñiz, J. (1992). *Teoría Clásica de los test*. España: Pirámide.
- Muñiz, J. (2001) *Nuevos desarrollos en la investigación psicométrica*. Facultad de Psicología Universidad de Oviedo. VIII conferencia Española de Biometría Pamplona España.
- Nunnally, C. J. (1970). *Introducción a la medición psicológica*. Argentina: Paidós
- Nunnally, C. J. (1991). *Teoría Psicométrica*. México: Trillas.
- Nunnally C. J. y Bernstein, I. (1995). *Teoría Psicométrica*. México: McGrawhill
- Pain, S. (1976). *Psicometría genética*. Buenos Aires: Nueva Visión.
- Pastor, M. y Sastre, R. S. (1994). Desarrollo de la inteligencia. En V. Bermejo (Ed.). *Desarrollo Cognitivo*. Madrid: Síntesis Psicológica.
- Pichot, P. (1979). *Los test mentales*. Buenos Aires: Paidós
- Pueyo, A. (1999) *Las inteligencias múltiples: la importancia de las diferencias individuales en el rendimiento*. Barcelona: Santillana.
- Rodríguez, J. (2001) *Psicometría*. *Revista electrónica de investigación educativa*. <http://www.psicometria.com>.
- Rojas, R. (1982). *Guía para realizar investigaciones sociales*. México: UNAM
- Roubertoux L. Y Cartier M (1996) ¿Es hereditario e CI?. *Mundo Científico*. Marzo, No. 166, pp. 264-273
- Sánchez, J. (2001) *Bassel, del sextante a la psicometría* <http://www.terra.es/personal/aasvega/articulo6.html>.
- Sanz, A. Lizarrago, M. L. y Sanz, R. (1998). La educación infantil y la inteligencia emocional. *Didáctica, comunidad educativa*. Febrero, pp.44-47
- Sanz, E. (2001) ¿Dónde mora la inteligencia? [www.cienciadigital.net](http://www.cienciadigital.net).
- Secadas, F. (1995). Inteligencia y Cognición. *Revista de Psicología General*, 48 (4). 511-537
- Serrano, M. (1984). *La medición en Psicología*. México: Trillas
- Silva, R. A. (1993) *Métodos Cuantitativos en Psicología. Un enfoque metodológico*. México: Trillas.
- Silva, R. A. (1998). *La investigación asistida por computadora*. México: UNAM
- Smith, M. (1971) *Estadística Simplificada para psicólogos y educadores*. México: Manual Moderno
- Sternberg, R. J. (1986). *Las capacidades humanas*. Barcelona: Labor Universitaria.
- Sternberg, R. J. (1990) *Más allá del cociente intelectual*. Bilbao: Biblioteca de Psicología.
- Sternberg, R.J. (1996). Countermyths, and Truths About Intelligence. En L. González V. *Teorías Científicas y Teorías Implícitas de la Inteligencia: Sus repercusiones en la escuela*. México: ITESM. Tesis Maestría en Educación.
- Sternberg, R. J. (1997) *Inteligencia Exitosa. Cómo una inteligencia práctica y creativa determina el éxito en la vida*. España: Paidós.

- Sternberg R. J. (1998) *Hacia mejores test de inteligencia*. En Wittrock M. C. y L. Baker. (Comp ) *Test y Cognición*. España: Paidós
- Vernon, P. (1980). *Inteligencia A, B y C*. Madrid: Ediciones Morova.
- Wechsler, D. (1978). *Escala de Inteligencia para adultos*. México: El Manual Moderno.
- Woolfolk, A. (1996). *Psicología Educativa* México. Prentice Hall.
- Yela, M. (1991) La estructura diferencial de la inteligencia. Enfoque factorial. En J. Mayor y J. L. Pinillos (Eds.) *Tratado de Psicología General*. Tomo V Madrid: Alambra.
- Zambrano. (1998). *La inteligencia y los test*. (mecanograma).