

227
2



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE PSICOLOGIA

"ESTUDIO DE CONFIABILIDAD Y VALIDEZ
DEL CORNELL INDEX."

TESIS

Que para obtener el Título de
LICENCIADO EN PSICOLOGIA

Presenta:

RUBEN WILLEBALDO VARELA DOMINGUEZ

Director de Tesis:

Lic. María Enedina Villegas Hernández



Junio de 1996.

TESIS CON
FALLA DE ORIGEN

TESIS CON
FALLA DE ORIGEN



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



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.

Este trabajo está dedicado a las siguientes personas:

A la memoria de mi madre, Socorro Domínguez,

A mi abuelita Virginia, mi tía Enedina y mi tía Virginia, gracias por su amor, cariño y generosidad, sin ustedes este trabajo no hubiese sido posible.

A mi padre, por ser quien es, por su cariño y apoyo.

A mis hermanos Rogelio, Miguel Angel, Guadalupe, Socorro y Germán

A todos ustedes mi cariño, y agradecimiento.

Agradecimientos

A la maestra Ma. Enedina Villegas Hernández, por su apoyo y paciencia, sobretodo por su ejemplo y enseñanzas a lo largo de mas de 10 años de trabajar a su lado.

A la maestra Josette Benavides Tourres, por su apoyo, cariño, y confianza,

Al comité de tesis integrado por Alma Mireia López , Samuel Jurado, Rocio Avendaño, mi agradecimiento y respeto.

Al Dr. Rogelio Díaz-Guerrero y su esposa la Dra. Lucy Rodríguez, por su apoyo, guía y confianza, y por permitirme el honor de trabajar a su lado.

De manera especial a mis amigos Angel Casales y Enrique Rivern, este trabajo llego a su término en gran parte gracias a ustedes., mi afecto y agradecimiento por siempre.

A Hilda Esquivel, Laura Chávez y Mónica Gutiérrez por su apoyo y amistad.

Al maestro Samuel Jurado Cárdenas por su amistad y entrega desinteresada.

A la maestra Rocio Avendaño Sandoval por ser una gran amiga y por sus valiosas enseñanzas.

Al maestro Humberto Zepeda, por sus enseñanzas y compañerismo.

A mis amigos Enrique, Pepe, Beto, Adolfo, por su confianza y amistad.

A mis compañeros de trabajo, Estela Cordero, Ricardo Díaz, Fabian Martínez, Enrique Buzo, Gerardo Hernández, Georgina Juárez por su amistad y apoyo desinteresado.

A todas aquellas personas, especialmente mis profesores, que directa o indirectamente contribuyeron a mi formación profesional.

A mis alumnos de la Facultad de Psicología, con los cuales diariamente el aprendizaje ha sido mutuo.

INDICE

RESUMEN.	1
INTRODUCCION.	2
Capítulo I. LA PERSONALIDAD	
1. La Personalidad.	4
1.1 La Definición.	4
1.2 La Clasificación de la Personalidad.	6
1.2.1 Teorías Dinámicas.	6
1.2.2 Teorías Personológicas.	6
1.2.3 Teorías Psicobiológicas.	7
1.2.4 Teorías Integradoras del Yo.	8
1.2.5 Teorías Factoriales.	10
1.3. La Evaluación de la Personalidad.	11
1.3.1 Historia de la Evaluación de la Personalidad.	11
1.3.2 Influencia de la medición psicológica.	15
1.3.3 Influencia de la Psicología Anormal.	18
1.3.4 Los Últimos Cincuenta Años.	20
1.3.5 Los Instrumentos Diferentes a las Pruebas.	24
1.3.6 Métodos Múltiples de Evaluación de la Personalidad.	25
1.3.7 Métodos y Definiciones sobre Evaluación de la Personalidad.	27
1.4. Evaluación Clínica: El Enfoque Empírico.	28
1.5. Descripción de la Personalidad y Evaluación Clínica.	30
1.6. Enfoques Teóricos de la Evaluación de la Personalidad.	31
1.7. Comparación de los Enfoques Teóricos y Empíricos.	32
1.8. Métodos de Evaluación de la Personalidad.	33
1.9 La Elección de los Procedimientos de Evaluación.	35
1.9.1 Características de los Procedimientos de Evaluación Estandarizados.	36
Capítulo II. TESTS PSICOLOGICOS	
2.1 Historia de la Medición en Psicología.	43
2.2 Definición de los Tests Psicológicos.	49
2.3 Los Reactivos o Ítemes que Conforman un Test.	50
2.4 Tipos de Reactivos o Ítemes.	51
2.5 Características de los Tests Psicológicos.	58

2.5.1 Grado de Dificultad y Poder Discriminativo de los Items.	58
2.5.2 Estandarización.	58
2.5.3 Confiabilidad.	59
2.5.3.1 Método de Confiabilidad Test-Retest.	61
2.5.3.2 Método de Confiabilidad con Formas Equivalentes o Paralelas.	61
2.5.3.3 Método de Confiabilidad de División por Mitades.	62
2.5.3.4 Método de Confiabilidad de Kuder-Richardson.	62
2.5.4 Validez.	62
2.5.4.1 Validez de Contenido.	63
2.5.4.2 Validez Empírica o de Criterio.	64
2.5.4.3 Validez de Construcción o Validez Estructural.	66
2.5.5 Normalización.	69
2.6 Clasificación de los Tests.	70
2.7 Influencia del Examinador en Cuanto a la Aplicación.	73
2.8 Puntos a Considerar sobre la Adaptación y Estandarización.	77
2.8.1 Problemas con los que debe enfrentarse el Traductor.	81

Capítulo III. EL CORNELL INDEX

3.1 Ficha Técnica.	87
3.2 Origen del Instrumento.	87
3.3 Propósito del Instrumento.	88
3.4 Confiabilidad.	91
3.5 Validez.	91
3.6 Administración.	91
3.7 Calificación.	92
3.8 Usos del Cornell Index.	93
3.9 Estudios en otros países.	94
TABLA A.	96
TABLA B.	97
TABLA C.	98
TABLA D.	99

Capítulo IV. METODO

4.1 Preguntas de Investigación.	100
4.2 Hipótesis de Investigación.	100
4.3 Variables.	100
4.4 Población.	100

4.5 Muestra.	100
4.6 Escenario.	101
4.7 Instrumento.	101
4.8 Diseño.	102
4.9 Tipo de Estudio.	102
4.10 Procedimiento.	103

Capítulo V. RESULTADOS, DISCUSION Y CONCLUSIONES

5.1 Análisis Estadístico.	104
5.2 Descripción de la muestra.	104
5.3 Poder discriminativo y Confiabilidad.	104
5.4 Validez.	116
TABLA 11.	117
TABLA 12.	118--120
5.5 Análisis de Resultados y Discusión.	121
5.5.1 Confiabilidad.	121
5.5.2 Validez.	122
5.6 Conclusiones.	124
5.7 Limitaciones.	126
5.8 Sugerencias.	127

REFERENCIAS DOCUMENTALES.	128
----------------------------------	-----

APENDICE.	133
------------------	-----

RESUMEN

El objetivo de esta investigación fue el obtener datos de confiabilidad y validez del cuestionario autodescriptivo de la personalidad que lleva por nombre Cornell Index al ser aplicado a una muestra de 765 estudiantes universitarios. Dado que no se contaban con estudios que apoyasen el uso de este instrumento, era necesario llevar a cabo esta investigación, ya que en nuestro país este y muchos otros instrumentos son usados de manera indiscriminada sin saber si son aplicables a nuestra cultura. Los resultados de la investigación mostraron lo siguiente: sólo 80 de los 101 reactivos que forman el Cornell Index mostraron ser sensibles para diferenciar entre sujetos que poseen padecimientos neuropsiquiátricos y psicósomáticos, de sujetos normales. Sin embargo, algunas de las áreas mostraron una confiabilidad muy baja y otras definitivamente fueron inconsistentes. Referente a la validez, se encontraron 10 factores que mostraron que la versión utilizada para México posee distinta estructura factorial a la versión original. Por tanto, se concluye que el instrumento llamado Cornell Index, en la versión FN2, no debe ser utilizado en población mexicana.

INTRODUCCION

El trabajo profesional de los psicólogos generalmente se ve apoyado por diversas herramientas, una de estas son los llamados tests psicológicos. Durante el proceso de formación del psicólogo este aprende, por un lado, los fundamentos de lo que es el diseño y la construcción de los llamados tests o pruebas psicológicas, y por otro las diversas aplicaciones que se pueden hacer de éstos en el campo de trabajo, sin embargo, el manejo de los tests se ve ensombrecido debido a que no hay un control adecuado sobre quien debe manejarlos, -desafortunadamente otros profesionistas hacen mal uso de estas herramientas- aunado a esto el 95% de los tests en nuestro país son de origen extranjero, de los cuales un porcentaje muy bajo cuentan con estudios serios y amplios que apoyen su uso para nuestra cultura. Esto trae como consecuencia el que se hagan evaluaciones diagnósticas que las más de las veces son falsas o erróneas, se toman decisiones con apoyo de instrumentos que no se sabe si son exactos (confiabilidad) o si miden lo que pretenden (validez).

Se torna un asunto muy delicado el utilizar instrumentos inadecuados, a partir de las cuales se va decidir el futuro de los seres humanos que evaluamos con ellos. La responsabilidad es demasiado grande e importante por un lado, por otro, la seriedad de nuestra profesión queda en entredicho, ya que esta pierde seriedad y credibilidad. En palabras de Dahlstrom (1993): *"las pruebas psicológicas pueden servir como un medio para equilibrar y minimizar el impacto de estas distorsiones únicamente si están firmemente construidas y se aplican responsablemente"*.

Esta investigación tuvo como objetivo el poder contar con datos de confiabilidad y validez que apoyasen el uso o no en población mexicana del instrumento llamado Cornell Index, en la versión FN2, ya que este forma parte del programa de prácticas de Evaluación de la Personalidad; el manual utilizado para este programa no cuenta con datos que permitan saber si el instrumento es adecuado o no para ser utilizado con nuestra población.

El primer capítulo de este reporte de tesis se centra en revisar todo lo referente al campo de la Personalidad, la definición, cómo se clasifica, y cómo se mide. El segundo capítulo ofrece un bosquejo general de todo lo referente a los Tests Psicológicos, en qué consisten, cómo se clasifican, qué características

deben de poseer, además de los principales puntos que deben de cubrirse cuando se analiza un instrumento de origen extranjero.

El tercer capítulo se aboca a presentar toda la información que fue posible recabar referente al Cornell Index, sus orígenes, propósitos, y características psicométricas que en la versión original lo caracterizaban; adicionalmente se habla de las investigaciones que en nuestro país se han llevado a cabo.

El cuarto y quinto capítulo se refieren al método que se siguió durante esta investigación, y los resultados encontrados para ésta.

Estudios como los de este tipo es indispensable que sean llevados a cabo para poder utilizar con confianza instrumentos creados originalmente para otras culturas, sin embargo, es importante señalar que en nuestro país es indispensable que se desarrollen instrumentos propios que tomen en cuenta las características idiosincráticas de la población, esto ayudará a contar con pruebas propias, por un lado, por otro, permitirá ser menos dependientes de instrumentos extranjeros que muy pocas veces o nunca toman en cuenta las características de nuestra cultura, y que origina que se siga dependiendo de enfoques en Psicología muchas veces ajenos a nosotros. Si esto no es posible, y si por una variedad de razones se emplean instrumentos contruidos para otra cultura, es necesario y obligatorio, por parte del psicólogo, investigar que estudios se han efectuado con dicho test o prueba en nuestro país, para garantizar que la evaluación que se realice sea lo más confiable y poder controlar, hasta donde sea posible los errores que se puedan cometer.

Capítulo I

LA PERSONALIDAD

1. LA PERSONALIDAD.

1.1 La Definición.

Un campo que siempre ha causado polémica dentro de la Psicología es el campo de la Personalidad. ¿Por qué?, porque toda persona y todo acontecimiento son algo único, ya que los hechos ocurren en el espacio y en el tiempo, y no pueden darse dos hechos en un mismo espacio y a un mismo tiempo. Sin embargo, entre las personas y entre los hechos puede darse la suficiente semejanza como para considerarlos idénticos.

En la Psicología, el campo de la personalidad se ocupa del individuo en cuanto a su totalidad y sus diferencias individuales. Aún cuando todas las personas son similares en algunos aspectos, el psicólogo se interesa por las diferencias existentes entre las personas, ¿Por qué unos tienen éxito y otros no?, ¿Por qué unos perciben las cosas de una manera y otros de otra?, ¿Por qué hay tanta diversidad de talentos?, o ¿Por qué unos enferman mentalmente y otros no?. El teórico de la personalidad intenta comprender la conducta a través de la interacción entre los distintos aspectos del funcionamiento individual.

Pero ¿Cómo definiríamos la personalidad?. En el momento actual no hay una definición absoluta y aceptada por todos. A nivel popular puede responder a dos acepciones. La primera, la identifica con la habilidad o la destreza social. La personalidad se evalúa por la eficacia para suscitar respuestas positivas en una cantidad de personas, en circunstancias diferentes; Ejemplo de ello son las famosas escuelas de "formación de personalidad", o bien, cuando un maestro se refiere a un alumno afirmando que sus habilidades sociales no son adecuadas para mantener relaciones satisfactorias con sus compañeros y sus maestros. La segunda considera que la personalidad del individuo está unida a la impresión más intensa o profunda que produce en los demás; así, se afirma que un individuo tiene una personalidad "agresiva" o que es muy "atrayente".

Ambas acepciones contienen elementos que encierran juicios de valor; según la descripción común las personalidades pueden ser buenas o malas. En la Psicología existen también problemas en referencia a esto, Allport encontró 50 definiciones diferentes que clasificó en categorías. Hace dos distinciones *definición biosocial* y *definición biofísica*. La primera corresponde al uso popular del término pues la equipara con un valor de estímulo social; la personalidad se define por la reacción de los otros. Y la segunda, afirma que la personalidad

radica en las características o cualidades del sujeto. En ella se indica la presencia de un aspecto orgánico en la personalidad, además del aspecto perceptible y supone que puede estar unida a las cualidades específicas del individuo, y que pueden ser medidas y descritas de manera objetiva.

Existe otro tipo de definición, la llamada *miscelánea*; en ésta, el término incluye cuanto atañe al individuo; el psicólogo elabora una lista de los conceptos que considera importantes para la descripción de la persona y manifiesta que la personalidad es la suma de todos ellos. Algunas otras definiciones dan mucha importancia a la función integradora y organizativa de la personalidad, la definen como *"aquello que da coherencia a los diversos tipos de conducta que manifiesta el hombre"*. Otras definiciones sugieren que la personalidad va unida a los aspectos individuales o peculiares de la conducta, el término es empleado para describir los rasgos distintivos de cada persona y que los diferencian de los demás.

Para concluir, otros teóricos consideran que la personalidad representa lo esencial del hombre.

El término alude a lo más representativo de cada uno. La definición de Allport ejemplifica este tipo de definiciones: *"la personalidad consiste en aquello que un hombre realmente es"*. Todas al final afirman que la personalidad es lo más típico y profundamente característico de la persona. (Hall y Lidzey, 1984; Engler, 1995).

El cómo definir la personalidad depende totalmente de la teoría a la cual se quiere atribuir particularmente. Así, **la personalidad esta definida por los conceptos empíricos particulares que integran la teoría de la personalidad empleada por el observador** (Hall y Lindzey, 1984). La personalidad consiste en un conjunto de valores o términos descriptivos que, de acuerdo con las variables o las dimensiones fundamentales de una teoría particular, se utilizan para definir la conducta de un individuo determinado. Ahora brevemente describiremos las principales teorías de la Personalidad. La manera en que están dispuestas u organizadas por los diversos autores, depende del criterio que han tomado para clasificarlas, ya sea por la época en que fueron elaboradas, por los métodos de trabajo o por la concepción teórica que apoya el autor. En el caso de este proyecto las describiremos en base a la sustentación teórica que apoyan:

1.2. LA CLASIFICACION DE LA PERSONALIDAD.

1.2.1. Teorías Dinámicas.

Conciben la personalidad como un conjunto de fuerzas que se oponen entre si, dando como resultado la conducta. Su principal representante es Freud, afirma que la conducta es producto de fuerzas opuestas en lucha por manifestarse, propone un aparato psíquico formado por 3 instancias: el ello, el yo y el super yo, estas tres instancias se contraponen y se encuentran en constante conflicto. Así mismo, describe el desarrollo humano por medio de etapas de resolución de conflicto psicosexual, que son de gran importancia para la estructuración de la personalidad (Hall y Lindzey, 1984; Engler, 1995). Ciertas situaciones de la vida diaria, ya sean externas o internas, pueden provocar fijaciones o regresiones en alguna de las etapas del desarrollo que pueden provocar alguna patología en particular, las etapas del desarrollo psicosexual son las siguientes:

ETAPA ORAL: Los labios y la cavidad bucal constituyen la primera zona erógena.

ETAPA ANAL: El ano es la principal zona erógena en esta etapa, es placentero para el niño el retener las heces fecales. Esta situada entre los 2 y los 4 años de edad.

ETAPA FALICA: Surge gracias a la unificación de las pulsiones de las dos etapas anteriores. Existe primacía por los órganos genitales como zona erógena.

PERIODO DE LATENCIA: Comprende la declinación de la sexualidad infantil antes del comienzo de la pubertad, el desarrollo psicosexual se encuentra detenido en esta etapa.

ETAPA GENITAL: Ultima etapa del desarrollo psicosexual caracterizada por la búsqueda de satisfacción sexual por medio de los órganos genitales.

1.2.2. Teorías Personológicas.

Su representante más importante es *Henry A. Murray*, su teoría es sintetizadora, incluye en ella al inconsciente, acentúa lo fisiológico, cubre aspectos sociológicos, cubre todo lo que atañe al hombre.

Dentro de estas teorías se le da gran importancia al concepto de "necesidad", la cual es definida como "un proceso hipotético cuya existencia se supone a fin de explicar algunos hechos, ciertos objetivos y otros subjetivos". (Murray, 1938; citado por Bischof, 1973). La necesidad puede originarse a partir de las funciones corporales internas, o bien, la situación que rodea a la persona.

Murray crea una taxonomía de necesidades que se utiliza mucho en la Psicología, estas son las siguientes: afiliación, agresión, autonomía, comprensión, consolación, contracción, defensa, deferencia, degradación, dominancia, evitación del dolor, evitación del fracaso, exhibición, juego, logro, orden protección, rechazo, sensibilidad y sexo (Geiwitz, 1977).

1.2.3. Teorías Psicobiológicas.

Sheldon y Kretschmer son los principales representantes de estas teorías, afirman que hay una estrecha relación entre el físico del individuo y su temperamento. Sheldon considera que hay tres componentes principales en la constitución física del hombre, éstos se mezclan de diferentes maneras en cada ser humano, son los siguientes:

ENDOMORFIA: Se refiere a la desigualdad entre el gran desarrollo de las vísceras con la consecuente acumulación de grasa, y el menor desarrollo de los huesos.

MESOMORFIA: Hay un predominio del desarrollo por los huesos y los músculos en oposición a otros componentes corporales.

ECTOMORFIA: Hay un mayor desarrollo de la piel y del tejido central en comparación a otros componentes.

Ahora bien, los componentes básicos del temperamento son los siguientes:

VISCEROTONIA: Sus rasgos principales son: el amor al confort, a la buena comida, y la inclinación a la vida social.

SOMATOTONIA: Hay tendencia a los riesgos físicos, agresividad competitiva, y la preferencia por el ejercicio físico.

CEREBROTONIA: Hay amor a la vida privada, retraimiento social y preferencia por actividades intelectuales.

Sheldon concluye que el temperamento se puede predecir teniendo como base la constitución corporal y viceversa.

Kretschmer propone a su vez tres tipos básicos de físico:

ASTENICO: Su físico es alargado, débil, hay estrechez en los hombros, extremidades largas y manos huesudas.

ATLETICO: Los brazos y las piernas son musculosas, complexión regular y hombros anchos.

PICNICO: Hay obesidad, posee grasa alrededor del cuello, tronco y particularmente el abdomen.

DISPLASTICO O MIXTO: Las diferentes partes del cuerpo no guardan proporción.

Cree que hay una relación entre estos cuatro tipos físicos con algunas psicosis, el esquizofrénico tiende a ser asténico, y el maniaco-depresivo, pícnico (Tyler, 1972).

1.2.4. Teorías Integradoras del Yo.

En este tipo de teorías se hace énfasis en el papel que tiene el Yo, como principal integrador de la personalidad. Sus principales representantes son *Rogers* y *Allport*.

Rogers elabora toda una corriente en psicoterapia llamada "*no directiva*", o "*centrada en el cliente*", este tipo de terapia sugiere que el ser humano tiene dentro de sí la capacidad y la motivación para curarse a sí mismo, el papel del terapeuta es únicamente de facilitador. Se define al Yo como una "configuración que tiende a aceptar los hechos que se conforman con el modelo anteriormente existente, y a rechazar los hechos que no se conforman". Da mucha importancia a la necesidad de actualización de necesidad de mirada positiva y empatía.

Allport define la personalidad como *"la organización dinámica dentro del individuo, de aquellos sistemas psicofísicos que determinan sus ajustes únicos a su ambiente"* (Allport, 1961; citado por Bischof, 1973). Este tipo de definición da la idea de hombre como una totalidad organizada y dinámica, se distinguen 5 niveles dentro de esta organización (Geiwitz, 1977):

- Respuesta condicionada.
- Hábitos (Sistemas integrados de respuestas condicionadas)
- Rasgos (Sistemas integrados de hábitos específicos)
- Yo (Sistemas de rasgos)
- Personalidad.

Aspecto importante dentro de la teoría de Allport, es el referente a la teoría de los rasgos que son definidos como *"un sistema neuropsíquico generalizado y orientado, que es peculiar en cada individuo y tiene capacidad para recoger un gran número de estímulos funcionalmente equivalentes y para iniciar y dar formas consistentes (equivalentes) de conducta adaptativa y expresiva"*. (Allport, 1961; citado por Bischof, 1973). En base a esto distingue 3 tipos de rasgos:

Rasgo Cardinal.- Es muy general, no hay pensamiento o acto que no se vea influido por el, es de naturaleza dominante.

Rasgos Centrales.- Son los que caracterizan al individuo.

Rasgos Secundarios.- Son los más limitados y fáciles de determinar, debido al tipo de respuestas que producen.

Allport considera que la inteligencia y el temperamento forman la materia prima de la personalidad, definiendo a esta como *"los fenómenos característicos de la naturaleza emocional de un individuo, fenómenos entre los cuales se encuentra su susceptibilidad a la estimulación, su intensidad y su velocidad de respuestas habituales, la cualidad de su estado de ánimo predominante y todas las peculiaridades de fluctuación e intensidad de su estado de ánimo; todos estos fenómenos dependen de la estructura, por tanto, son de origen hereditario"* (Allport, 1961; citado por Bischof, 1973).

1.2.5. Teorías Factoriales.

Los principales representantes de estas teorías son *Cattell* y *Eysenck*, ellos le dan gran importancia al uso de las matemáticas en el estudio de la personalidad, ya que creen que las matemáticas expresan las relaciones entre variables de una manera más precisa. En especial, ellos utilizan el Análisis Factorial como herramienta en sus investigaciones.

Cattell define la personalidad como "aquello que permite predecir lo que hará una persona en una situación determinada"; (*Cattell*, 1950, pag. 2; citado por *Bischof*, 1973). Así, aquello que permite predecir la conducta lo define como Rasgos Originales, éstos se obtienen como resultado de aplicar el Análisis Factorial a grupos de datos relacionados con la vida cotidiana. Hace una clasificación sobre los diferentes tipos de rasgos:

Rasgos de Molde Ambiental: se originan a partir de la sociedad y la cultura en que vivimos, se aprenden.

Rasgos Constitucionales: son de origen orgánico o interno, pueden deberse a la herencia, y pueden cambiar con el desarrollo del individuo.

Además hace otra distinción:

Rasgos Dinámicos: tienen siempre una meta, la conducta varía según los incentivos que va encontrando.

Rasgos de Aptitud: Estos implican una conducta que varía según la complejidad de la tarea.

Rasgos de Temperamento: Estos no son afectados por los incentivos, ni por la complejidad en las tareas a realizar.

Una aportación importante de *Cattell*, es la referente a los rasgos dinámicos y su supuesto origen, los rasgos dinámicos constitucionales los define como *ergios*, y los rasgos dinámicos de origen ambiental los llama *engramas*. A partir de la definición de estos tipos de rasgos *Cattell* elabora una ecuación de especificación, que permite predecir la respuesta de un individuo según el grado en que posea cada rasgo original (*Geiwitz*, 1977).

Eysenck afirma que la personalidad es *"la organización íntegra de todas las características cognoscitivas, afectivas, conativas y físicas del individuo que se manifiestan en directo enfoque con otros"*. (Eysenck, 1953; citado por Bischof, 1973; Pervin, 1993). Sugiere que se utilice el Análisis Factorial de discriminación múltiple para lograr identificar tipologías; así, logra establecer tres dimensiones primarias de la personalidad:

INTROVERSION (Super yo)	-- EXTRAVERSION (Ello)
NEUROTISISMO	-- NO NEUROTISISMO
PSICOTISISMO	-- NO PSICOTISISMO

Es muy común que se dé un cambio de patrón a otro.

1.3. LA EVALUACION DE LA PERSONALIDAD.

1.3.1. Historia de la Evaluación de la Personalidad.

La conducta interpersonal del hombre no es solamente de interés para sus amigos y vecinos sino también de relevancia social general, y el interés en la evaluación de la personalidad data de mucho tiempo. Si un hombre prehistórico tratara de dilucidar las intenciones de un visitante extraño en su campamento al observar su postura, esto podría considerarse como un esfuerzo primitivo de evaluación de la personalidad. El empleo posterior de adivinos y oráculos para determinar la veracidad de las ofertas de paz del enemigo podría considerarse de modo semejante. Así, hay muchos procedimientos antiguos e informales que anteceden al desarrollo de la psicología moderna, guardando un valor histórico y a que aclaran algunas de las cuestiones metodológicas involucradas en la evaluación de la personalidad y que en cierto grado se siguen practicando en varias sociedades, a pesar de su desarrollo científico y de su posición contemporánea bastante dudosa, como por ejemplo: la astrología, la quiromancia y la frenología, métodos que son considerablemente más antiguos que cualquier técnica actual de evaluación profesional (Lanyon, 1977; Aiken 1995, 1996; Gregory, 1993).

ASTROLOGIA.

La astrología, intentó prever los acontecimientos de la Tierra a través de la observación de las estrellas fijas y otros cuerpos celestes, se cree que se haya originado hace cerca de veinticinco siglos en la Mesopotamia. La creencia de que las estrellas eran dioses poderosos llevó a los antiguos a concluir que los sucesos humanos podrían ser predichos por el estudio de los cielos. La personalidad de cada individuo, y la causa de los acontecimientos de su vida, eran determinadas por su horóscopo (la determinación de las estrellas al momento de su nacimiento), es decir, la evaluación de la personalidad de un individuo se llevaba a cabo considerando el momento de su nacimiento y obteniendo después la información predictiva apropiada de uno de varios manuales o almanaques completos.

Si bien no existe evidencia que muestre que el momento del nacimiento determina la personalidad, ha habido cierto interés por, y apoyo para, la propuesta de que la "estación" de nacimiento tiene algunas consecuencias modestas en la conducta posterior. Algunos estudios han mostrado que las personas nacidas durante los meses del verano tienen inteligencia ligeramente superior a las nacidas en el invierno (Pintner y Forlano, 1943; Orme, 1965, citados por Lanyon 1977 y Aiken 1996), aunque también ha habido hallazgos no concluyentes o en sentido opuesto (Davies, 1964; Craddick, 1966, citados por Lanyon, 1977). La explicación más razonable para los descubrimientos positivos parecer ser la de que el calor extremo de los meses del verano puede afectar adversamente al feto en desarrollo, dando por resultado un nivel más bajo del funcionamiento cerebral más tarde en la vida para los niños gestados en los meses del verano y nacidos en el invierno. No obstante, aún se carece de evidencia definitiva en el área.

Así, la cuestión de una relación entre el momento del nacimiento y la personalidad, problema con raíces en la astrología, se ha vuelto a exponer de maneras que la hacen susceptible a la investigación con la metodología de una ciencia empírica.

QUIROMANCIA

La Quiromancia se refiere a la determinación de las características de un individuo por la interpretación de las diversas irregularidades y pliegues de la piel de la palma de la mano. De la quiromancia se sabe que existía como sistema generalizado en China alrededor de 3000 años antes de Cristo, aunque su inicio y raíces teóricas se pierden en la antigüedad. La quiromancia ofrece una evaluación

completa de la personalidad de los individuos mediante la lectura de signos como: líneas de las manos, protuberancias o **montículos** entre estas líneas. La ausencia total de cualquier explicación razonable de las inferencias interpretativas, junto con la clara noción de que los montículos y otras características de la mano pueden cambiar con el ejercicio físico, condujeron al rechazo de la quiromancia como superstición y charlatanería. El quiromántico en su práctica formula declaraciones triviales generalizadas que podrían aplicarse a cualquiera, a lo cual Meehl (1956, citado por Lanyon 1977 y Aiken 1996) sugirió la expresión "efecto Barnum" para señalar "aquellos procedimientos clínicos seudo exitosos en los que se hacen descripciones del paciente a partir de las pruebas de modo que encajen con el paciente **general o totalmente en virtud de su trivialidad**; y en los que cualesquiera inferencias **no triviales**, pero quizá erróneas, se ocultan en un contexto de afirmaciones o negaciones que logran gran confianza sencillamente a causa de las tasas base de la población". Tallent (1958), citado por Lanyon (1977) denominó a esta clase de evaluación de la personalidad como portador del error de **Tía Fanny**, porque contiene **más que nada información** que sería cierta para la "Tía Fanny" de cualquiera. El grado en que cualquier procedimiento de la evaluación de la personalidad involucre el efecto Barnum o de "Tía Fanny" por que produzca la descripción trivial, altamente generalizada de la personalidad, siempre tiene que ser cuidadosamente valorada.

FRENOLOGIA

La frenología, arte de la evaluación de la personalidad por medio de la medición de la forma **externa del cráneo humano**, recibió sus principales impulsos de Franz Joseph Gall, un médico y anatomista alemán, ya avanzado el siglo dieciocho. La suposición básica de Gall era que el cerebro humano constituía el **centro de control de la conducta humana**, un punto de vista que ahora es aceptado uniformemente por los psicólogos. Las operaciones de la corteza cerebral, eran muy poco comprendidas y la opinión científica de la época se pronunciaba por una **localización estricta de la acción cortical**, en la que cada función o facultad se centraba en una región definida y **específica** de la superficie cerebral. El tamaño de estas regiones creaba las alteraciones correspondientes y observables en el **tamaño del cráneo**, lo cual podía ser utilizado por el frenólogo para evaluar las diversas características del individuo, incluyendo la personalidad. Por ejemplo, una prominencia del cráneo en el área de la honestidad indicaría un individuo honesto. Así, la frenología habilitó a sus practicantes para dar una evaluación completa de la personalidad basada, en este caso, en las protuberancias y contornos del cráneo. Las suposiciones básicas concernientes a la especificidad

del funcionamiento cerebral y a la formación del cráneo ahora han sido claramente refutadas, y por consiguiente la teoría sobre la cual descansa la frenología está desacreditada.

Gall, trataba de relacionar la conducta con el funcionamiento cerebral y la forma del cráneo a través de la observación personal. El examinaba los cráneos de personas vivas y luego trataba de estudiar sus cerebros por medio de la autopsia después de la muerte. Asimismo examinaba los cráneos de personas que se encontraban en hospitales psiquiátricos, prisiones, universidades y otros lugares donde pudieran encontrarse individuos con deficiencias o dotes excepcionales, esperando poder relacionar la forma del cráneo con las características que habían reunido a estas personas.

Paradójicamente, no fue la falta de validez científica de la postura de Gall lo que dio lugar a las objeciones más serias. Más bien las impugnaciones más enérgicas se basaron en fundamentos filosóficos (Davies, 1955, citado por Lanyon 1977). Los críticos de la frenología insistían en que la postura de Gall no era indemostrada sino indemostrable; y aun si fuera demostrable, era inmoral. La confirmación de que la frenología, según estos críticos, implicaría una aceptación del ateísmo, el fatalismo y una negación de la responsabilidad moral (los mismos tópicos que se plantean en la objeción a los puntos de vista deterministas contemporáneos de la conducta humana). La frenología perdió así su carácter de ciencia experimental, si bien basada en suposiciones inadecuadas, convirtiéndose en un culto de índole religiosa, cerrado a cualquier auto examen o modificación que se basara en datos nuevos.

La utilización de signos físicos estables y la búsqueda de relaciones nítidas entre los atributos físicos de un individuo y sus características psicológicas ha sido continua, remanente de la interminable preocupación del hombre acerca de la relación entre la mente y el cuerpo. Esta preocupación tuvo una expresión anterior en la teoría humoral, una fisiología antigua que se mantuvo en boga a través de la Edad Media. La teoría proponía que el cuerpo contenía cuatro humores o fluidos principales: sangre, flema, bilis amarilla (cólera) y bilis negra (melancolía). La proposición particular de estos cuatro ingredientes en personas diferentes determinaba sus "complexiones" o características de personalidad, sus cualidades físicas y mentales, sus disposiciones particulares. Así se sostenía que era posible, mediante la lectura de los signos físicos de las cantidades relativas de estos humores en un individuo dado, la determinación de su personalidad distintiva.

La suposición de que la personalidad del hombre es un reflejo de su cuerpo físico se encuentra también en los trabajos modernos de Sheldon y sus colegas (Sheldon, Stevens y Tucker, 1940, citados por Aiken, 1996). La teoría de Sheldon de los Somatotipos divide a las personas en tres tipos principales: el *ectomorfo* que es delgado y frágil, el *mesomorfo* que es poderoso y muscular, y el *endomorfo* que es redondo, blando y obeso. Cada uno de estos tipos ideales tiene una personalidad específica que es consecuencia innata de esta particular estructura corporal. Las características de la personalidad de la mayoría de los individuos, que son combinaciones de estos tres tipos de cuerpo, son determinadas por las proporciones relativas de estos factores de tipo corporal en su somatotipo individual. La teoría de Sheldon ha sido tema de controversia, y la evidencia de su validez es, en el mejor de los casos equívoca.

No puede negarse que hay relaciones significativas entre ciertos signos físicos y algunas características de la personalidad, sin embargo los intentos sin éxito a lo largo de la historia por apoyarse en signos físicos parece tener por lo menos una de las siguientes dos características: primera, en apoyarse en una teoría que postulaba que las características de la personalidad eran "causadas" por atributos físicos particulares; y segunda, el carecer de una conexión lógica entre el signo físico y el atributo de la personalidad.

No obstante que la frenología es ya un intento por tratar de estudiar y evaluar a la personalidad de una forma más experimental, esta, la astrología y la quiromancia, junto con métodos tales como la consulta de la bola de cristal, la lectura en las hojas de té y la fisiognomía (evaluación de la personalidad por medio de los rasgos faciales), no han sobrevivido a las investigaciones científicas sobre su utilidad, y la práctica de tales procedimientos dudosos se considera charlatanería en nuestros días.

1.3. 2. Influencia de la medición psicológica.

La evaluación científica de la personalidad tiene sus raíces en el estudio de las "diferencias individuales" por medio de la medición psicológica. La mayoría de los estudiosos de las diferencias individuales ubican la apertura de esta área de indagación en 1796, cuando un ayudante de un astrónomo de Greenwich, Kinnebrook, fue despedido porque había observado constantemente los tiempos de tránsitos estelares casi un segundo más tarde que su superior. Posteriormente se cayó en cuenta que esta diferencia en "tiempo de reacción" era estable, y que

las personas diferían entre sí en esta y otras características mensurables de una manera similar estable.

El estudio de las diferencias individuales recibió considerable impulso gracias a la obra de Darwin sobre la evolución. A fin de estudiar los efectos de la genética humana, sería necesario identificar claramente las diferencias individuales en aquellas conductas que tuvieron valor de adaptabilidad y de supervivencia en la especie humana. Sir Francis Galton, un famoso sabio británico del siglo diecinueve, se interesó por la herencia de estas diferencias. Aunque había habido intentos anteriores por categorizar y medir lo que Galton llamó "*facultades intelectuales*", él mismo fue el iniciador de la medición de las facultades no intelectuales, conocidas más comúnmente en ese tiempo como "*carácter y temperamento*".

Galton expuso la medición de varias de tales facultades en 1884. Su declaración "*que el carácter que modela nuestra conducta es un algo definitivo y duradero, y por ello... es razonable tratar de medirlo*", tiene una marcada semejanza con la mayoría de los puntos de vista contemporáneos de la personalidad. Más aún, Galton propuso que la medición del carácter debe involucrar el mismo enfoque "*... que yace en la raíz de los exámenes de la capacidad intelectual*", y expuso sucintamente los métodos para la medición de la emoción y el temperamento, utilizando lo que pudiéramos denominar el enfoque de la muestra de la conducta. Por ejemplo, sugirió medir la emoción por medio de instrumentos que registrarán los cambios fisiológicos en el individuo durante situaciones de tensión de la vida real, sugiriendo tomar el pulso y medir los cambios en el volumen del muslo con la contracción arterial, medidas todas que se usan todavía en los laboratorios de investigación psicofisiológica para estudiar las reacciones emocionales. Respecto al temperamento, Galton estaba convencido de que podrían desarrollarse técnicas similares para la evaluación del buen y mal genio en los humanos. (Lanyon, 1977; Aiken, 1996).

El estudio de las diferencias individuales en E.U.A. fue iniciado por James Mckeen Cattell que trabajó por un corto tiempo con Galton en Inglaterra, para luego regresar a los Estados Unidos en 1888 estableciendo el Laboratorio Psicológico en la Universidad de Pensilvania. Si bien sus intereses se centraban en el área de la psicofísica, la percepción y el tiempo de reacción, tuvo una fuerte influencia sobre el desarrollo de otros instrumentos de medición psicológica, incluyendo las pruebas de personalidad.

Más o menos al mismo tiempo, en Francia, Alfred Binet, inició una serie de estudios de personas eminentes en las artes y las ciencias. Binet utilizó una serie normalizada de **tareas experimentales**, incluyendo la observación de tipos corporales, mediciones de la cabeza y caligrafía. Inicia así mismo una serie de investigaciones sobre el funcionamiento mental, que incluían a la personalidad, empleando una **amplia variedad** de respuestas tales como el conocimiento de palabras, el razonamiento y la capacidad numérica. Estas investigaciones llevaron al desarrollo de las ahora famosas pruebas Binet de Inteligencia. Algunas de sus tareas, que implicaban el relato de cuentos acerca de grabados y la identificación de manchas de tinta, fueron antecedentes de las pruebas proyectivas de la personalidad.

De esta manera, antes de 1915, la investigación sobre la medición de la personalidad, fue precedida de trabajos sobre la medición de las destrezas o las capacidades, y surgió del antiguo interés académico por la medición de las diferencias individuales humanas. Galton había introducido el uso de muestras directas de conducta en situaciones de la vida real, despertando considerable interés tanto en E.U. y Francia. Dos de los seguidores de Galton, Karl Pearson y Charles Spearman, estaban desempeñando un papel importante en el desarrollo de procedimientos estadísticos que aportaron técnicas poderosas para el trabajo posterior de la evaluación.

Por ese tiempo, un científico inglés de nombre Webb publicó lo que parece haber sido el primer intento considerable por resumir los aspectos importantes del carácter a través del estudio intensivo de un gran número de sujetos. Obtuvo valoraciones de un grupo grande de hombres en edad escolar y universitaria sobre 40 o más cualidades, las cuales tenían "*... un efecto general y fundamental en la personalidad total*". Las valoraciones se obtuvieron de personas que conocían bien a los sujetos, y los datos resultantes fueron tratados mediante una forma inicial de los procedimientos estadísticos que ahora se conocen como análisis factorial. Su interpretación indicó que el aspecto más básico de la personalidad era la "*...consistencia de acción resultante de la volición deliberada o voluntad*". Dicho estudio es digno de tomarse en cuenta tanto por su uso de las valoraciones de jueces como por sus técnicas estadísticas refinadas (Lanyon, 1977).

Weeb también resumió dos estudios previos que habían enfocado el estudio científico de la personalidad de una manera diferente, mediante el uso de lo que llamó el "*método biográfico*", un análisis de autobiografías escritas. Estos pueden

considerarse los precursores del enfoque actual del "*formulario de datos biográficos*" para la evaluación de la personalidad.

1.3.3. Influencia de la Psicología Anormal.

Por el mismo tiempo al interés en el estudio de las diferencias individuales, otros investigadores se estaban dedicando a la evaluación formal de la conducta humana **anormal**. Este interés surgió de la necesidad práctica de clasificar y categorizar las diversas clases de psicopatología, así como de la necesidad más teórica de lograr una mejor comprensión de los fenómenos en escrutinio. Estrechamente relacionados con estos esfuerzos fueron los intentos paralelos de Binet por elaborar pruebas clínicas que identificarían a los niños que carecían de la capacidad mental para aprovechar la educación escolar, y que aportarían información teórica acerca de sus limitaciones intelectuales.

Se considera en general que los primeros intentos de evaluación de personalidad de casos psicopatológicos implicaron los procedimientos de **asociación de palabras**. Emil Kraepelin, renombrado psiquiatra alemán que fue responsable primordial de nuestros procedimientos actuales de clasificación psiquiátrica, hizo cierto uso de una "*prueba de asociación libre*" ya en 1892, y Sommer, uno de sus colegas, sugirió su utilización para diferenciar varias clases de perturbación mental en 1894 (Anastasi, 1988). En la técnica de asociación de palabras, al sujeto se le presenta una palabra estímulo, frecuentemente de una lista estandarizada de tales palabras, y se le pide que responda tan rápidamente como sea posible con la primera palabra que le venga a la mente. Se hace un registro de las respuestas dadas y del tiempo de reacción a cada presentación.

El uso de la asociación de palabras como un método para la identificación de los conflictos inconscientes de personalidad fue propuesto inicialmente por Carl Jung. Jung (1910), citado por Aiken, (1996) elaboró una lista estandarizada de palabras que fueron consideradas especialmente útiles en la detección de áreas de conflicto o perturbación, indicadas por los tiempos de reacción excesivamente largos, el no responder, los malentendidos de las palabras estímulo, y otras reacciones emocionales, como tartamudear o sonrojarse.

Otro enfoque más formal empleado por Kent y Rosanoff (1910, citados por Lanyon, 1977) sobre la asociación de palabras consiste en una lista estandarizada que omitía deliberadamente palabras que eran especialmente susceptibles de evocar experiencias personales. Desarrollaron normas de respuestas comunes, al

considerarse los precursores del enfoque actual del "formulario de datos biográficos" para la evaluación de la personalidad.

1.3.3. Influencia de la Psicología Anormal.

Por el mismo tiempo al interés en el estudio de las diferencias individuales, otros investigadores se estaban dedicando a la evaluación formal de la conducta humana **anormal**. Este interés surgió de la necesidad práctica de clasificar y categorizar las diversas clases de psicopatología, así como de la **necesidad** más teórica de lograr una mejor comprensión de los fenómenos en escrutinio. Estrechamente relacionados con estos esfuerzos fueron los intentos paralelos de Binet por elaborar pruebas clínicas que identificarían a los niños que carecían de la capacidad mental para aprovechar la educación escolar, y que aportarían información teórica acerca de sus limitaciones intelectuales.

Se considera en general que los primeros intentos de evaluación de personalidad de casos psicopatológicos implicaron los procedimientos de **asociación de palabras**. Emil Kraepelin, renombrado psiquiatra alemán que fue responsable primordial de nuestros procedimientos actuales de clasificación psiquiátrica, hizo cierto uso de una "*prueba de asociación libre*" ya en 1892, y Sommer, uno de sus colegas, sugirió su utilización para diferenciar varias clases de perturbación mental en 1894 (Anastasi, 1988). En la técnica de asociación de palabras, al sujeto se le presenta una palabra estímulo, frecuentemente de una lista estandarizada de tales palabras, y se le pide que responda tan rápidamente como sea posible con la primera palabra que le venga a la mente. Se hace un registro de las respuestas dadas y del tiempo de reacción a cada presentación.

El uso de la asociación de palabras como un método para la identificación de los conflictos inconscientes de personalidad fue propuesto inicialmente por Carl Jung. Jung (1910), citado por Aiken, (1996) elaboró una lista estandarizada de palabras que fueron consideradas especialmente útiles en la detección de áreas de conflicto o perturbación, indicadas por los tiempos de reacción excesivamente largos, el no responder, los malentendidos de las palabras estímulo, y otras reacciones emocionales, como tartamudear o sonrojarse.

Otro enfoque más formal empleado por Kent y Rosanoff (1910, citados por Lanyon, 1977) sobre la asociación de palabras consiste en una lista estandarizada que omitía deliberadamente palabras que eran especialmente susceptibles de evocar experiencias personales. Desarrollaron normas de respuestas comunes, al

aplicar la lista a 1000 personas normales. Al comparar las respuestas de los normales con la de los psicóticos, pudieron demostrar que los psicóticos daban un número impresionantemente menor de respuestas comunes que los normales y que estas diferencias eran lo suficientemente grandes para la significación diagnóstica en los casos individuales. De tal suerte, la prueba de asociación de palabras puede ser considerada el primer instrumento psicométrico práctico para identificar a las personas emocionalmente perturbadas, aunque hoy día es poco utilizada con este propósito.

Una preocupación similar por los problemas de la adaptación llevó a Heymans y Wiersma en 1906 a elaborar una lista de síntomas indicadores de psicopatología. Esta lista revisada por Hoch y Amsden en 1913 y más tarde por Wells en 1914, influyó en la creación del inventario de personalidad por autoreferencia. El primero de tales inventarios fue la (Woodworth's Personal Data Sheet) **Hoja de Datos Personales** de Woodworth, que fue desarrollada durante la Primera Guerra Mundial para diagnosticar la capacidad de los soldados para adaptarse satisfactoriamente a las tensiones y restricciones de la vida militar. Un Comité de Adecuación Emocional fue designado por el Consejo Nacional de Investigación para trabajar a cargo de Woodworth, siendo su solución un inventario a papel y lápiz, titulado la Hoja de Datos Personales para atenuar la desconfianza acerca de su verdadero propósito, que era una entrevista psiquiátrica estandarizada en forma impresa (Aiken, 1996). Las preguntas se basaban en síntomas neuróticos comunes y en los informes de síntomas que eran observados realmente en hombres que no habían sido capaces de adaptarse a la guerra y a sus tensiones. La puntuación en este cuestionario era el número de reactivos respondidos en el sentido considerado típico de inadaptación o psiconeurosis. Posteriormente a este cuestionario se le llamó Inventario Psiconeurótico de Woodworth (Woodworth Psychoneurotic Inventory) que fue el precursor de las múltiples pruebas de personalidad de papel y lápiz actualmente en uso.

Otro desarrollo digno de tomarse en cuenta en la evaluación de la personalidad con propósitos clínicos fue elaborado por Hermann Rorschach, un psiquiatra suizo, que había estado utilizando la percepción de manchas de tinta como un enfoque de los problemas teóricos de la psicología cuando hizo el descubrimiento empírico de que los resultados podían aprovecharse para hacer diagnósticos psiquiátricos diferenciales. Su obra principal fue publicada en alemán en 1921, pero no apareció oficialmente en inglés hasta 1942. Si bien investigadores anteriores se habían valido de las manchas de tinta como material de asociación libre, se había limitado a un análisis de contenido temático. La

contribución de Rorschach fue el intentar un análisis sistemático del interés del sujeto por los aspectos **formales** de las manchas tales como el color, las sombras, el movimiento aparente y si el sujeto prefería considerar la mancha como un todo o en partes. El creía que estos distintos enfoques perceptuales estaban relacionados con diferentes procesos o "estructuras" psicológicos de la personalidad, siguiendo una orientación empírica y estadística. La obra de Rorschach ha sido semilla, tanto para estimular la investigación con sus manchas de tinta como para el desarrollo de otras situaciones estímulo **minimamente** estructuradas, para los cuales pueden provocarse respuestas y analizarlas de acuerdo con el propósito de la evaluación de la personalidad.(Lanyon, 1977).

1.3.4. Los últimos cincuenta años.

Puede considerarse que los primeros días o comienzos de la evaluación contemporánea de la personalidad abarcaron hasta el año de 1920, haciéndose ostensibles dos líneas claras de desarrollo, influenciadas claramente por el estudio académico de las diferencias individuales y la evaluación clínica de la psicopatología, además de otras influencias adicionales.

Quizá la influencia adicional más importante fue la de la teoría psicoanalítica, que supuso que las causas de la conducta estaban "*enterradas muy por debajo del conocimiento consciente*". La comprensión más significativa y útil de la personalidad de un individuo debe basarse en un análisis de estos aspectos "más profundos" de su funcionamiento, más que en la conducta **manifiesta** que se encuentra fácilmente disponible al escrutinio superficial. Importante asimismo fue la influencia de la psicología de la Gestalt, con su implicación de que la personalidad del individuo completo era más que la simple suma de rasgos conductuales aislados.

Estos intereses e influencias condujeron al clínico a medir la **personalidad** sobre una preocupación por los aspectos ocultos o encubiertos de la personalidad, y llevaron a la popularización de lo que Frank (1939, citado por Aiken 1996) denominó posteriormente **técnicas proyectivas**. Los métodos proyectivos fueron conceptuados como aquellos que presentan al sujeto una situación para la cual hay pocos patrones culturales de respuestas claramente definidos, de modo que él debe "proyectar" sobre ese campo ambiguo "*su manera de ver la vida, sus entendimientos, significaciones, patrones y especialmente sus sentimientos*". Según Frank, una prueba proyectiva comporta la presencia de una situación estímulo que desencadena el significado y la organización característicos y

privados, del mundo íntimo del individuo que él ha traído a la situación de examen y el cual es llevado a entrar en juego en respuesta a las demandas de la prueba. Frank consideró estos métodos como indirectos y que abordan el patrón de organización interna y estructura de la personalidad sin desintegrar o modificar tal patrón como existe, muy a la manera de una radiografía del cuerpo. (Aiken, 1996).

Aparte de las manchas de tinta de Rorschach que halló preferencia inmediata entre los clínicos, especialmente aquellos de orientación psicoanalítica, apareció otro grupo de métodos proyectivos, que Lindzey (1961), citado por Lanyon, (1977) ha llamado técnicas constructivas fue desarrollado por Henry A. Murray y sus colegas de la Universidad de Harvard en 1930 (Morgan y Murray, 1935, Murray, 1938, citados por Lanyon 1977). Murray, con un fuerte interés por la creación literaria, elaboró la Prueba de Apercepción Temática (Thematic Apperception Test) (TAT), en la cual los sujetos crean relatos acerca de grabados ambiguos. De acuerdo con Murray, el contenido de los relatos refleja las experiencias pasadas y necesidades presentes del sujeto.

Un enfoque constructivo similar se encuentra en la Prueba de Cuatro Grabados (Four Picture Test) de Van Lennep (1951), citado por Lanyon (1977), desarrollada inicialmente por los años treinta, y en instrumentos posteriores tales como los grabados de Blacky (Blacky Pictures) (Blum, (1950) citado por Lanyon (1977) y la Prueba de Apercepción para Niños (Children's Apperception Test) (Bellak, 1954, citado por Lanyon, 1977). Otro grupo de procedimientos proyectivos ampliamente utilizados involucra un análisis de los propios dibujos y pinturas creativas del sujeto. Si bien los dibujos de la figura de un hombre habían sido utilizados por algún tiempo como un método conveniente para estimar la inteligencia (Goodenough, 1926, citado por Lanyon, 1977), correspondió a Machover (1949, citados por Lanyon 1977) y a Buck (1948, citado por Lanyon, 1977) popularizarlos como un método proyectivo para la evaluación de la personalidad.

La segunda tendencia en la medición de la personalidad (diferencias individuales), es ejemplificada por el trabajo sobre inventarios de personalidad de papel y lápiz. Dichos instrumentos podían ser aplicados con facilidad a muchos individuos en un grupo, la puntuación era rápida y objetiva, y podían emplearse procedimientos estadísticos para determinar las normas, la consistencia interna, la relación entre las puntuaciones de la prueba y otras medidas conductuales, la estructura factorial y otros aspectos del instrumento. Este enfoque hizo énfasis en

la medición de rasgos o aspectos aislados más que en una evaluación global de la personalidad total. Los reactivos de prueba contenidos en estos instrumentos a menudo tenían que ver con los aspectos más superficiales y obvios de la conducta manifiesta y su susceptibilidad al engaño deliberado planteó muchos problemas para la interpretación de las puntuaciones obtenidas.

Siguiendo la obra original de Woodworth, hubo varios intentos bastante inmediatos por adaptar su cuestionario a otros grupos (niños de escuela, delinquentes juveniles y estudiantes universitarios), esencialmente mediante la modificación del fraseo de las preguntas. Una variante del procedimiento de cuestionario fue desarrollada por Pressey (1919, citado por Lanyon 1977), cuya Prueba X-O (X-O Test o "Cross out" cuya abreviatura fonética es X-O) presentaba al sujeto listas de palabras en lugar de preguntas. Al sujeto se le pedía que leyera la lista y tachara con una cruz las palabras que considerara malas, que lo pusieran nervioso o ansioso, o que involucraban sus gustos o intereses.

El contenido de los reactivos de estos primeros inventarios era elegido por su **validez aparente**: esto es, los reactivos eran incluidos si el elaborador de la prueba pensaba que suscitarían respuestas diferentes de los sujetos bien adaptados y mal adaptados. El Inventario de Ajustes de Bell (Bell Adjustment Inventory) (Bell, 1939, citado por Lanyon 1977) es un ejemplo algo más reciente de este enfoque de selección de reactivos y construcción de pruebas.

Una técnica diferente y más refinada de selección de reactivos es el **enfoque empírico**, en el que las respuestas reales de diferentes grupos de sujetos son examinadas a fin de determinar cuáles respuestas son características de ese grupo. El enfoque empírico, usado por Rorschach en el desarrollo original de su procedimiento psicodiagnóstico, fue empleado por Edward K. Strong (1927, citado por Lanyon 1977 y Aiken, 1996) en la elaboración del Cuestionario de Intereses Vocacionales de Strong (Strong Vocational Interest Blank), presentando un importante modelo inicial de dicho enfoque en la construcción de pruebas. La técnica empírica estuvo igualmente involucrada en cierto grado en la construcción del Inventario de Personalidad de Bernreuter (Bernreuter Personality Inventory) y en la Escala de Temperamento de Humm-Wadsworth (Humm-Wadsworth Temperament Scale). Esta última fue predecesora directa del más cuidadosamente elaborado Inventario Multifásico de la Personalidad de Minnesota (Minnesota Multiphasic Personality Inventory, MMPI) (Hathaway y Mckinley, 1940; 1951, citados por Aiken 1996).

Uno de los problemas difíciles implicados en el uso de los inventarios de personalidad de papel y lápiz fue la tendencia de algunos sujetos a evitar el dar respuestas socialmente indeseables a los reactivos, en vez de dar las respuestas más apropiadas personalmente. El MMPI trató de afrontar este problema incluyendo varias **escalas de validez**, como la escala de mentira, que da cierta medida de la tendencia a responder de una manera socialmente aceptable. La preocupación por este problema también llevó al desarrollo, en los años 40's, de la técnica de elección forzada, que fue utilizada posteriormente por Edwards (1953, citado por Lanyon 1977) en su **Cédula de Preferencias Personales** (Personal Preference Schedule). Dado que el procedimiento de elección forzada requiere que el sujeto acepte uno de dos reactivos de igual deseabilidad (o no deseabilidad) social, se supuso que su respuesta reflejaría algo más que la aceptabilidad social obvia de los reactivos aislados.

Goldberg (1970, citado por Lanyon 1977), al comentar la proliferación actual de inventarios de personalidad, hizo una distinción entre dos clases de inventarios, basados en las razones de su desarrollo. El primer grupo de inventarios fue elaborado en respuesta a las presiones de la sociedad para que se trataran problemas aplicados específicos como los de ajuste personal, satisfacción y éxito en la elección vocacional y los que tocaban aspectos de aprovechamiento académico más allá de lo que es predecible a partir de las medidas de la aptitud escolar o la inteligencia. El segundo grupo de inventarios se basó en concepciones de la estructura de las diferencias individuales más que en cualesquiera consideraciones sociales o de la vida real y se consideró como surgido de conceptos teóricos acerca de la naturaleza de la personalidad, como los inventarios de introversión-extroversión y de masculinidad-feminidad y también las medidas tipo inventarios basadas en dos teorías importantes sobre las diferencias individuales, la clasificación de "valores personales" de Spranger (1928, citado por Lanyon 1977) y el esquema para el ordenamiento de las necesidades manifiestas de Murray (1938, citado por Lanyon 1977). Aunque el esquema de clasificación de Goldberg fue ideado solamente para medidas tipo inventario de la personalidad, tiene clara relevancia para otros métodos de evaluación de la personalidad y tiende a separar las dos influencias históricas principales sobre la metodología de la evaluación.

Al paso de los años ha habido una tendencia persistente por parte de muchos seguidores de cada una de las dos tradiciones -las técnicas proyectivas y los inventarios de personalidad- a tener un fuerte prejuicio emocional negativo contra el trabajo de la otra tradición. Esta tendencia probablemente se acrecentó

en 1939 al denominar Frank a la primera corriente como "técnicas proyectivas" con la declaración explícita de que estos instrumentos proporcionaban medidas indirectas de la estructura de la personalidad, y la implicación de que los entonces más tradicionales inventarios no estaban afrontando la verdadera substancia de la personalidad, trayendo como consecuencia la pérdida de mucho intercambio clínico y científico entre ambos campos. De esta forma con la posible excepción de la técnica de Manchas de Tinta de Holtzman (Holtzman's Inkblot Technique), han sido insignificantes los adelantos de la década pasada que representarían una síntesis de las dos tradiciones. (Gregory, 1993).

Se ha señalado que los inventarios de autoreferencia pueden ser, y muy a menudo son, contestados en forma proyectiva, y que los "perfiles" de personalidad resultantes son utilizados por clínicos como base para una interpretación global de la personalidad, en términos que involucran bastante más que una simple descripción de la conducta manifiesta. Por otro lado, las técnicas proyectivas tradicionales pueden ser calificadas objetivamente y luego interpretadas de acuerdo con datos normativos y evidencia empírica en relación con el significado de ciertas categorías de respuestas aisladas; por lo que dicha distinción es espuria y que no sirve para un propósito útil.

1.3.5. Los Instrumentos diferentes a las pruebas.

Otros enfoques formales en la evaluación de la personalidad, son las formas de datos biográficos y las técnicas de muestra de la conducta. La forma de datos biográficos, ampliamente utilizada con propósitos de evaluación en situaciones industriales y de negocios, ha tenido poco uso sobre una base práctica para los propósitos clínicos, con la excepción de un interés por ciertos signos predictivos o pronósticos obtenibles a través de instrumentos como la escala de Phillips (1953, citado por Aiken, 1996) para predecir los resultados de la esquizofrenia. Dicha escala es un método para estimar el ajuste premórbido de los pacientes psiquiátricos al asignar una ponderación diferencial a diferentes reactivos en la historia clínica del paciente, principalmente a aquellas conductas que entrañan intereses y relaciones heterosexuales. Los datos biográficos también han sido utilizados a lo largo de mucho años en la elaboración de "tablas base de expectativas" para predecir si un recluso es o no susceptible de violar su libertad condicional si le es otorgada; a pesar de esto, no hay un instrumento clínico de este tipo utilizado ampliamente dada la importancia por los mismos trabajadores profesionales de salud mental a la información biográfica.

La técnica de muestra de conducta también han sido sólo un uso limitado en la evaluación rutinaria de la personalidad. Notables excepciones han sido los estudios de Hartshorne y May (1928, citado por Aiken, 1996) sobre la personalidad y el proyecto de evaluación clínica de la Oficina de Servicios Estratégicos de E.U.A. durante la Segunda Guerra Mundial.

Más recientemente, el muestreo de conductas ha sido utilizado en la investigación sobre terapia conductual (Lang y Lazovik, 1963, citado por Aiken, 1996) para la evaluación de ansiedades específicas.

1.3.6. Métodos múltiples de evaluación de la personalidad.

Los métodos múltiples de evaluación hacen referencia al desarrollo de los procedimientos para la evaluación comprensiva por un grupo o jurado de evaluadores profesionalmente preparados. En dicho procedimiento los individuos que se van a evaluar son reunidos en grupos pequeños para un estudio de varios días en un "centro de evaluación" que suele ser un lugar residencial remoto. Los propósitos de este tipo de evaluación pueden ser los de entender mejor a individuos de este tipo, como arquitectos, psicólogos o clínicos creativos, o seleccionar candidatos para algún puesto difícil de alto nivel, como astronautas o agentes secretos. Por supuesto, estos procedimientos también son útiles para entender en general los tópicos implicados en la personalidad y su evaluación. Taft ha proporcionado un resumen ilustrativo de diversos intentos por elaborar métodos múltiples de evaluación de la personalidad.

Muchas clases de datos pueden ser puestos a disposición del jurado de evaluadores. Típicamente estarían incluidos una variedad de datos biográficos, evaluaciones intelectuales, respuestas a cuestionarios de autoreferencia y técnicas proyectivas, desarrollo de varias entrevistas, y una diversidad de otras muestras conductuales o pruebas situacionales que serían elaboradas por su relevancia para el grupo particular bajo escrutinio. Entre los diversos problemas inherentes a este enfoque están el cómo deberían ponderarse los múltiples índices conductuales (puntuaciones de pruebas, valoraciones y otros similares), cómo se han de combinar en un solo predictor, y cómo han de desarrollarse las medidas del criterio (Aiken, 1996).

El primer programa de evaluación múltiple parece haber sido desarrollado por Henry Murray (1938, citado por Lanyon 1977 y Aiken, 1996) y sus colaboradores en Harvard y consistía principalmente en evaluaciones

comprendidas y detalladas de una serie de personas con el objeto de entender mejor el funcionamiento de las personas normales. Muchas de estas técnicas de Murray fueron empleadas para seleccionar espías durante la Segunda Guerra Mundial por E.U.A. e Inglaterra. Durante 1946 y 1951 la Administración de Veteranos (Veterans Administration) y la Universidad de Michigan (Kelly y Goldberg, 1959, citados por Lanyon, 1977 y Aiken, 1996), emplearon un programa extenso de evaluación múltiple para mejorar la selección de estudiantes prometedores para la profesión de psicología clínica y Holt y Luborsky (1958, citados por Lanyon, 1977) lo llevaron a cabo pero con residentes psiquiátricos como sujetos.

Con un propósito algo diferente pero valiéndose de metodología esencialmente similar, el grupo de psicólogos del Instituto de Evaluación e Investigación de la Personalidad de la Universidad de California (IPAR) ha investigado las características de personalidad de individuos altamente creativos y exitosos en una serie de disciplinas. Sus informes de investigación (Mackinnon, 1961, citado por Lanyon, 1977) tienden a dar fuerte apoyo a la utilidad de este enfoque de centros de evaluación cuando involucra muestras conductuales relevantes.

En general podría resumirse que la cantidad de métodos de evaluación como la frenología, la astrología y la quiromancia son fascinantes pero no se considera que tengan la utilidad documentada y requerida para su utilización seria y profesional por los psicólogos especializados en la evaluación de la personalidad.

En referencia a los instrumentos de evaluación reconocidos profesionalmente, pueden verse dos corrientes históricas en su desarrollo. La primera es rastreable hasta el interés académico original en la medición de las diferencias individuales, y tiene sus consecuencias más importantes en los inventarios contemporáneos de papel y lápiz y las pruebas de personalidad de autoreferencia. La segunda se originó ante la necesidad por comprender la conducta clínicamente anormal, y se considera como conducente al interés por y al desarrollo de las técnicas proyectivas.

1.3.7. Métodos y Definiciones sobre Evaluación de la Personalidad.

Después de haber revisado la historia de como ha surgido el interés por el estudio de la evaluación de la personalidad, es necesario entender los conceptos y definiciones sobre este constructo para entender mejor su surgimiento y desarrollo hasta nuestras fechas.

En general, los psicólogos profesionales evalúan la personalidad humana por dos razones principales: primera y primordial, para ayudar a nuestra comprensión de la conducta de una persona particular, esto es, la evaluación clínica de un individuo solo; y segunda, en el contexto de la investigación psicológica y la elaboración de la teoría, para hacer avanzar nuestro conocimiento general de la conducta humana.

El uso principal de los instrumentos de evaluación de la personalidad en la psicología clínica y de personal se ha dado en el primer contexto: la evaluación clínica de la persona sola. Aquí, la meta es llegar a alguna decisión acerca del curso de acción futuro de un individuo determinado, o hacer una predicción acerca de su conducta futura singular. Este uso clínico o profesional es el que ha constituido la razón principal del desarrollo y existencia continuada de la mayoría de los instrumentos de evaluación mejor conocidos.

Respecto del segundo uso de la evaluación, los investigadores de la psicología emplean las técnicas de evaluación de la personalidad por varias razones. Algunos desean estudiar un concepto particular de la personalidad como la rigidez, o la fuerza del yo, o la necesidad de logro, ya sea por sí mismo o dentro del esquema de algún enfoque teórico de la personalidad. Otra razón podría hallarse en el contexto de lograr una mejor comprensión de un cierto síndrome clínico o patrón de conducta. Una tercera razón es la de reducir la gama de las diferencias individuales que existe entre un grupo de sujetos en una tarea de investigación para seleccionar a los sujetos conforme a una característica distintiva de la personalidad, por ejemplo estudiar si los sujetos con alta y baja ansiedad difieren en cuanto a su capacidad para relajarse, lo que puede afectar en la enseñanza de la relajación para ambos grupos. En los usos investigativos de la evaluación, el interés recae típicamente en las diferencias de grupo más que en un individuo particular o en la comprensión de las implicaciones de los hallazgos para la persona sola. El investigador está interesado primordialmente en cómo los resultados generales arrojaron luz sobre algún hecho general o ley conductual que se refleje en la conducta de las personas de su experimento.

Así, el desarrollo de la evaluación de la personalidad se ha centrado en la preocupación por resolver problemas prácticos aplicados o individuales y en el interés por una mejor comprensión teórica a nivel general del funcionamiento humano.

1.4. EVALUACION CLINICA: EL ENFOQUE EMPIRICO.

Supongamos que tenemos evidencia de que una respuesta particular o conjunto de respuestas (una puntuación de un inventario de personalidad de autoreferencia o conjunto de respuestas a una prueba proyectiva) tiene alta correlación con alguna conducta no evaluada en la prueba (el criterio). Pudiera ser que se supiera que las puntuaciones altas en la prueba ABC se correlacionan altamente con el éxito en las ventas de puerta en puerta. En el enfoque empírico de la evaluación clínica, el psicólogo necesitaría solamente aplicar cualquiera de estas pruebas, comparar los resultados obtenidos con los hallazgos de investigación empíricos, y hacer su predicción en cuanto a la probabilidad de éxito o fracaso del individuo por ejemplo en las ventas. El psicólogo por lo común no conoce, ni le concierne, la conexión psicológica o teórica entre las respuestas del individuo a la prueba y la conducta a predecir. Aquí no es necesario entender ¿por qué? hay una conexión entre las puntuaciones altas en la prueba ABC y la conducta del vendedor; es suficiente con saber que tal relación existe en forma confiable. Naturalmente, es importante conocer las condiciones limitantes de esta relación derivada empíricamente; esto es, conocer los grupos particulares de individuos, las clases de ventas para los que vale esta relación y las circunstancias en las que pudiera fallar. Pero, si suponemos que dicha información está disponible y que el individuo es extraído del segmento de la población cubierto por una relación empírica dada, podemos valernos de este conocimiento para la predicción o la toma de decisión acerca de ese individuo.

El enfoque empírico podría representarse esquemáticamente como sigue:

<p>Información procedimientos de — evaluación.</p>	<p>— empíricos —</p>	<p>Hallazgos empíricos — previos.</p>	<p>Decisión acerca del individuo.</p>
---	----------------------	--	--

Dependiendo de factores como la magnitud de la relación entre el predictor y el criterio, la representatividad de los datos empíricos originales y la similitud del criterio actual con el original, el psicólogo profesional estará más o menos

acertado en su predicción o decisión, siempre y cuando la probabilidad de error grave no es grande. No obstante, otros teóricos (Peak 1953, Loevinger 1957, citados por Lanyon, 1977) y otros han sostenido que un enfoque puramente empírico, con una confianza total en las relaciones observadas, es limitado y superficial porque el psicólogo no será capaz de desarrollar noción alguna, salvo a un nivel correlacional primitivo, de ¿por qué? la suya fue una decisión útil (Lanyon, 1977).

El enfoque empírico de la evaluación implica un supuesto básico. Este es el supuesto de la consistencia de la conducta interpersonal, es decir, que hay patrones duraderos de respuestas conductuales que trascienden a las situaciones y que pueden ser abordados y usados para entender a los individuos particulares y a las personas en general. La búsqueda de una comprensión de estas características duraderas de las personas sirve para identificar el campo integro de la personalidad como una área de indagación. Si bien una preocupación por la naturaleza de estas características duraderas es elaramente el interés de los enfoques de la evaluación de la personalidad basados más teóricamente, el enfoque empírico debe asimismo dar por centrada su existencia. Ello es particularmente importante para entender este supuesto, porque una serie de psicólogos han empezado a cuestionar la existencia de las características interpersonales duraderas. Esta corriente teórica reciente conceptúa a la personalidad en términos de conducta social aprendida compleja, utilizando la terminología de la teoría del aprendizaje conductual general. Este enfoque representado por los escritos de Kanfer y Saslow (1965), Wallace (1966), Peterson (1968), y Mischel (1968, 1969) citados por Lanyon (1977), niega que las tendencias disposicionales generalizadas tengan efectos importantes sobre la conducta, poniendo mayor énfasis en las determinantes ambientales actuales de las conductas relacionadas con la personalidad, y en las recompensas y castigos relevantes que estén operando sobre estas conductas. En este punto de vista, la comprensión y predicción de la conducta interpersonal de un individuo se haría no tanto a partir de una evaluación de su "personalidad subyacente" como de un análisis de sus respuestas manifiestas y de los contornos ambientales en que estas respuestas fueron dadas (Lanyon, 1977; Aiken, 1996; Goldstein, 1984).

1.5. DESCRIPCIÓN DE LA PERSONALIDAD Y EVALUACIÓN CLÍNICA.

En el enfoque empírico de la evaluación de la personalidad el psicólogo profesional se desplaza directamente de los datos psicométricos obtenidos a la predicción o decisión sin considerar cualesquiera procesos intermedios. Este enfoque se basa en datos y su uso se limita a aquellas situaciones en que se ha demostrado que los datos psicométricos son de valor predictivo. Si bien los datos que tenemos pueden predecir el éxito de ventas de puerta en puerta, también estamos interesados en el potencial para las ventas en una gran variedad completa de situaciones y podríamos afirmar que todo lo que se requiere en estas circunstancias es ampliar el enfoque empírico a estas nuevas situaciones, es decir, hacer acopio de nuevos datos empíricos adicionales en cada situación específica de interés. Desafortunadamente esto no sucede en la práctica. Más bien, los datos existentes son interpretados en términos más generales, utilizando el lenguaje no específico de la descripción de la personalidad, y luego son empleados para la predicción o la toma de decisiones. Así, las personas que obtienen puntuaciones altas en las pruebas ABC llegan a ser caracterizadas como alta en "ventas". El enfoque de descripción de la personalidad respecto a la evaluación pudiera representarse esquemáticamente como sigue:

Información de los procedimientos de — evaluación.	Descripción de la personalidad —	Decisión acerca del individuo
---	---	--

Así, hay muy poca "teoría" involucrada como no sea el sentido común y la elección de los conceptos implicados en esta clase de procedimientos está determinada primordialmente por consideraciones pragmáticas más que teóricas.

Conviene precisar lo que se entiende por teoría: Marx (1963), identifica tres significados generales de "teoría" en los escritos psicológicos contemporáneos. El uso más general está en referencia a cualquier proceso conceptual en la ciencia, en contraste con los datos empíricos u observacionales. Un segundo uso se hace en referencia a cualquier principio explicativo generalizado, principio que ordinariamente entraña una declaración de relaciones funcionales entre variables. Un tercer se refiere a un grupo de leyes organizadas lógicamente o relacionadas deductivamente, basadas típicamente en observaciones empíricas. Este último uso se apega más estrechamente a los de las ciencias más desarrolladas y también tiene una relación más estrecha con el

concepto de sistema, en el sentido de que este último término se refiere ampliamente dentro de la psicología a un agrupamiento de proposiciones teóricas e inclinaciones metodológicas.

Así parecería que los índices desarrollados con propósitos predictivos o de toma de decisiones son relativamente específicos para esa situación y que resultan mucho menos eficientes cuando la situación a predecir difiere de cualquier forma substancial, siendo siempre necesario hacer una comprobación empírica sobre la generalidad de los datos predictivos, ya que si no verificamos empíricamente la generalidad de nuestras conclusiones, corremos el riesgo de hacer predicciones incorrectas y de tomar malas decisiones y ha veces lo haremos peor que si nos hubiésemos atendido a la pura casualidad al hacer nuestras elecciones. El problema principal que entraña el empleo de tales constructos descriptivos como "ventas" es que su vaguedad representa una clara invitación al abuso por parte de muchos usuarios potenciales (Lanyon, 1977; Goldstein, 1984).

El procedimiento de utilizar las relaciones empíricas entre puntuaciones de las pruebas y las conductas no evaluadas por la prueba para apoyar la utilidad de un constructo abstracto generalizado, introduce el importante concepto de **validez de constructo** (Cronbach y Meehl, 1955, citado por Aiken 1996): *El supuesto de que una puntuación o índice de una prueba es indicativo de alguna característica interna generalizada del respondiente en casos en que hay muy pocos datos disponibles que apoyen tal supuesto.*

1.6. ENFOQUES TEORICOS DE LA EVALUACION DE LA PERSONALIDAD

La información es procesada y una descripción de la personalidad del individuo es provista por el psicólogo, formulada en el lenguaje de su teoría favorita de la personalidad. Por ejemplo, supongamos que un individuo que se considera necesitado de psicoterapia es remitido a un psicólogo, de inclinación freudiana, para una evaluación de la personalidad. Luego de completar los pasos que abarcan una evaluación clínica, el psicólogo expide su informe, incluyendo su hallazgo de que el paciente tiene un conflicto de Edipo irresuelto. El psicólogo recomienda inclusive que el paciente sea visto en terapia por un terapeuta hombre dominante, como arreglo óptimo para la resolución del conflicto. Debe ser claramente patente que la conclusión de que el paciente requiere un terapeuta hombre no se desprende directamente de los datos de la evaluación, sino más bien

de algunas consideraciones teóricas por parte del psicólogo. Este proceso pudiera expresarse como sigue:

Información proveniente de los proce- dimientos de evaluación	-----	Descripción de la personalidad	-----	Teoría de la personalidad	-----	Decisión acerca del paciente
--	-------	---	-------	--	-------	---

Para que ésta sea una decisión útil para el paciente, tres condiciones deben mantenerse ciertas:

- a) la interpretación por el psicólogo de los datos de evaluación debe ser correcta dentro del esquema de la teoría particular de la personalidad que se esté usando;
- b) su comprensión de la teoría debe ser lo bastante adecuada para permitirle tomar una decisión que sea consistente con las demandas de la teoría y,
- c) la teoría misma debe ser útil.

El éxito o fracaso con el paciente dependerá de las condiciones antes mencionadas, de su combinación y de las circunstancias que envuelvan la vida del paciente y la del terapeuta. El éxito con un individuo aislado no valida esta bastante larga cadena de inferencia; por una razón, los mismos factores extraños (circunstancias favorables de la vida real del paciente o competencia extrema del terapeuta) pueden haber estado operando.

La validez de este enfoque de la evaluación individual no puede determinarse adecuadamente observando los resultados de las decisiones basadas en él, sino que debe investigarse a través de un cuidadoso estudio paso a paso de la cadena inferencial. En otras palabras, validar la técnica de evaluación está estrechamente ligado con el validar la teoría de la personalidad utilizada en la cadena (Lanyon, 1977; Pervin, 1993).

1.7. COMPARACION DE LOS ENFOQUES TEORICOS Y EMPIRICOS.

Dado el estado actual de la teoría de la personalidad, el psicólogo clínico tiene una mayor probabilidad de acertar en su toma de decisiones si se vale del enfoque empírico más bien que del teórico. Hay menos pasos e inferencias entre

los datos y la decisión en el enfoque empírico, porque es algo ya probado y que se puede volver a probar, lo cual reduce la probabilidad de error (Lanyon, 1977; Aiken, 1996).

No es posible sostener que el enfoque empírico no es la evaluación de la personalidad en lo absoluto, sin embargo, han estado tradicionalmente tan ligados a la evaluación de la personalidad, particularmente en el contexto de establecer la " validez " de las pruebas de personalidad que deben ser considerados parte integral del campo.

Así, las teorías existentes de la personalidad no son lo suficientemente explícitas para ofrecer superioridad alguna a la predicción a través del enfoque teórico. Por otro lado, este enfoque ofrece más para el avance último del campo, de modo que su desarrollo se hará más lento si los procedimientos empíricos son favorecidos siempre para el trabajo aplicado.

Es posible en principio, desarrollar directamente un instrumento de evaluación a partir de una teoría de la personalidad. En dicho caso, las respuestas del individuo inmediatamente tendrían sentido dentro del contexto de la teoría y no sería necesaria la actividad inferencial adicional alguna por parte del clínico. Uno de los pocos instrumentos que están claramente y explícitamente basados en la teoría es la Prueba de Grabados de Blacky (Blacky Pictures Test) (Blum, 1949) (citados por Lanyon, 1977).

1.8. MÉTODOS DE EVALUACION DE LA PERSONALIDAD.

El término "evaluación de la personalidad" se refiere al proceso de recoger y organizar la información acerca de otra persona con la expectativa de que esta información conducirá a una mejor comprensión de la persona (Lanyon, 1977 y Aiken 1996). Comprender a otro individuo y a su personalidad comporta típicamente el hacer alguna predicción acerca de la conducta futura de esa persona en términos bastante específicos. Entre las clases de predicciones que han de hacerse podrían estar su probabilidad de éxito vocacional en una variedad de áreas, su respuesta a diferentes clases de tratamientos psicológicos o su probabilidad de demostrar ciertas conductas socialmente indeseables. La tarea del psicólogo profesional entraña a menudo participar en una variedad de decisiones que implican a otras personas, basándose en la evaluación de sus personalidades.

Sin embargo, todos nosotros nos vemos continuamente implicados en la recolección de una diversidad de información no sistemática acerca de la conducta de las personas que nos rodean y de continuo estamos, si no completamente conscientes, haciendo inferencias acerca de sus características de personalidad a partir de nuestras observaciones. Expresiones como "*Lo conozco perfectamente bien*" o "*Lo he conocido por mucho tiempo pero no lo entiendo en lo absoluto*", reflejan nuestras maneras cotidianas de hablar sobre nuestro conocimiento de la personalidad de otro.

Los procesos informales o **intuitivos** por los que llegamos a conocer y entender a las personas con las que estamos estrechamente involucrados son similares a los empleados por los psicoterapeutas para entender a su paciente. El enfoque informal se caracteriza por la gran cantidad de tiempo que implica y también por la falta de especificidad tanto en la forma de recolección de datos como en la del análisis de estos. Se espera que cada observador proceda con la tarea de **entender** a la otra persona a su propio ritmo y a su propia inespecificada manera.

Aunque el enfoque intuitivo **resulta** fascinante, no es considerado como un procedimiento sistemático de afrontar la tarea de evaluación. Entre los **métodos** utilizados por el psicólogo profesional en su papel formal como evaluador experto de la personalidad se encuentran: Los instrumentos de evaluación estandarizados de forma sistemática y la entrevista (una técnica de evaluación no estandarizada).

Los instrumentos de evaluación estandarizados se pueden identificar en tres grupos; el primero consiste en las que se consideran generalmente como **pruebas de personalidad** (**Inventarios** y técnicas proyectivas). Un segundo grupo de instrumentos hace uso de información factual específica acerca del estado actual del individuo y de su historial pasado, la clase de información que pudiera obtenerse de un cuestionario de datos biográficos o de una **entrevista** cuidadosamente estructurada. Este método, el **método de los datos biográficos**, es ampliamente utilizado en la selección de personal en la **industria** y un tercer grupo de instrumentos proporciona modelos **miniatura de situación de la vida real**, en los cuales puede observarse directamente la conducta de interés. Por ejemplo, un instrumento para evaluar la tolerancia a la tensión del **sujeto** consistiría en una situación estructurada y estandarizada en la que se requeriría del individuo que se **desempeñara** bajo cierta tensión producida artificialmente. La situación se arreglaría normalmente de modo que la conducta pudiera ser observada

directamente sin que la persona se percatara de lo que estuviera ocurriendo. Este método de evaluación se conoce como **muestras de conducta**.

Algunos psicólogos han tenido la impresión de que es más apropiado considerar a ciertas pruebas de personalidad, particularmente las técnicas proyectivas, como situaciones estandarizadas para suscitar una muestra de conducta. Ellos piensan que la situación de examen proyectivo representa un modelo estandarizado de una interacción interpersonal con demandas ambiguas y que la situación integra -los materiales de la prueba proyectiva, más el examinador, más sus instrucciones, más el ambiente inmediato- constituye el "instrumento de evaluación". (Lanyon, 1977; Aiken, 1996).

1.9. LA ELECCION DE LOS PROCEDIMIENTOS DE EVALUACION.

El psicólogo al hacer una evaluación de la personalidad se vale tradicionalmente de una entrevista, pruebas, datos biográficos y otros datos de historial como valoraciones supervisadas o calificaciones escolares y cualesquiera otras herramientas evaluativas adicionales que pueda escoger, a fin de proveerse de información acerca del individuo en estudio. Aun cuando los pasos involucrados en la evaluación clínica normal puedan ser más organizados que los métodos de evaluación informal intuitivo, no hay procedimientos normalizados para reunir los datos que sean seguidos universalmente por los psicólogos y los instrumentos de evaluación escogidos por un psicólogo particular estarán en función de su preparación y experiencia, sus preferencias personales y su familiaridad con la literatura actual (Lanyon, 1977; Aiken, 1996).

Un problema recurrente en la elección de los métodos de evaluación es el de decidir entre las pruebas de personalidad formales y los métodos intuitivos informales. Suponiendo que los dos métodos arrojarían predicciones igualmente correctas, aunque no suele ser el caso, se tendría que tomar en cuenta el "factor tiempo", "los factores económicos", "la eficiencia práctica de alguno de los métodos" y "las ganancias relativas o redituación de los cursos alternativos de acción", entre otros factores.

Puede haber ocasiones en que no sea justificable aplicar ningún procedimiento de evaluación en lo absoluto. Tomando en consideración todos los factores, si las pérdidas (en tiempo, gastos, hostilidad incrementada del paciente, pérdida de la moral del personal) sobrepasan a las ganancias (información pertinente adicional sobre el individuo, probabilidad incrementada de lograr la

predicción o decisión óptimas), entonces sería mejor no utilizar procedimiento de evaluación alguno. Desafortunadamente, ha habido poca investigación sistemática sobre las ganancias relativas de las técnicas de evaluación. Hathaway (1959, citado por Lanyon 1977) consideró el tiempo gastado por los psicólogos clínicos en la evaluación clínica de rutina y estimó que si se hiciera una evaluación empírica de su eficiencia, se abandonarían más de 40 % de estas actividades.

Hay una gran variedad de métodos de evaluación de la personalidad: métodos informales o intuitivos, entrevistas, procedimientos estandarizados formales, datos biográficos y muestras conductuales. A continuación se describirán de manera breve algunos de estos métodos y sus características tomando en cuenta el enfoque sobre el que están elaborados:

1.9.1 Características de los procedimientos de evaluación estandarizados.

Los instrumentos de evaluación estandarizados satisfacen las siguientes condiciones:

- 1) *los estímulos que se usan en el proceso de evaluación son idénticos para todos los sujetos a evaluar, y siempre son presentados de la misma forma.*
- 2) *existen normas disponibles o distribuciones de frecuencia de las respuestas, ya sea formales o informales (intuitivas), de modo que las respuestas puedan ser ubicadas en un lugar específico dentro de una gama anticipada, y*
- 3) *hay correlatos útiles de personalidad y conductuales de las respuestas a observar.*

Las condiciones anteriores sugieren tres dimensiones que podrían ser útiles para caracterizar los procedimientos evaluativos de la personalidad. La primera dimensión está relacionada con la variedad de respuestas relevantes dadas por el sujeto a examinar y que son observadas por el que examina y que pueden denominarse *grado de estructura de la respuesta*. Un alto grado de estructura o estandarización del procedimiento es una condición importante para cualquier instrumento formal de evaluación. De tal manera, que por ejemplo los procedimientos para aplicar el Rorschach deben ser tan inflexibles como los de la aplicación del MMPI. La mayor parte de los inventarios de personalidad de lápiz y papel, como son el MMPI o el CPI, son considerados como altamente estructurados, ya que toda vez que las únicas respuestas que serán registradas son

las de "verdadero" y "falso" a una serie de declaraciones personales. Cualquier otro tipo de respuesta no está tomada en cuenta por las instrucciones preliminares y son ignoradas si son hechas a pesar de las instrucciones. Como el MMPI o el CPI se resuelven marcando una hoja de respuestas o bien una hoja óptica para calificación automatizada, las variaciones en la conducta como son el vigor o el tiempo de respuesta no son registradas, y también el procedimiento de calificación de ambos instrumentos permite el uso de semejantes datos.

Por el contrario, la prueba de manchas de tinta de Rorschach es considerada como relativamente inestructurada ya que el examinador observa y registra una amplia gama de conductas del examinado, y las considera como material de respuesta relevante. Dentro de esto queda contemplado como la forma, el color o el sombreado, el contenido real de la respuesta, cambios en el tono, volumen, intensidad y rapidez del habla, etc.

Una segunda dimensión en la que difieren los procedimientos de evaluación es el *grado* en que se hallan disponibles las normas formales publicadas. Los inventarios estructurados, por ejemplo el MMPI o el CPI, con una gama menor de respuestas relevantes, se prestan más de por sí al desarrollo de normas formales. Estas son elaboradas por los autores, o en su defecto por el investigador que está adaptando el instrumento a otra cultura. Son presentadas a través de tablas en donde a partir de puntajes brutos son transformadas en calificaciones estandarizadas para facilitar su posterior interpretación. Por otro lado, las técnicas proyectivas que proporcionan una gama mayor de respuestas para ser consideradas, generalmente no cuentan con normas formales disponibles. En este caso el psicólogo que use estos procedimientos se vale de su experiencia para poder interpretar las respuesta que le son dadas por el examinado. Existen algunos intentos por tratar de presentar algún tipo de normas en la calificación de estos procedimientos como por ejemplo la calificación del Rorschach en el sistema de Exner (Aiken, 1996), y que generalmente no son aceptados en su totalidad.

Una tercera dimensión es la *utilidad* del instrumento. Esta se puede entender por el grado en que las respuestas del examinado permitirán la comprensión y predicción de algunas de sus conductas no evaluadas por las pruebas. Un instrumento muy útil permitiría que fueran hechas predicciones relativamente precisas acerca de conductas socialmente importantes, como el éxito en una ocupación particular o el pronóstico en psicoterapia. Los instrumentos menos útiles son menos precisos en sus predicciones o se refieren a

la conducta menos significativa, como el desempeño en una tarea experimental de laboratorio.

La posición de un instrumento de evaluación en cada una de estas tres dimensiones --*grado de estructura de la respuesta, disponibilidad de normas y utilidad*-- está establecido por un aspecto básico del instrumento, y se refiere al método de construcción del mismo, particularmente al método empleado para seleccionar los reactivos. Por lo general hay tres enfoques o estrategias principales en la construcción de los instrumentos de evaluación formales (Aiken, 1996; Kline, 1983, 1993; Lanyon 1977):

1) *el Teórico-Racional*

2) *el Empírico*

3) *el de Consistencia Interna*

En el *enfoque Teórico-Racional* un instrumento puede estar basado sobre consideraciones puramente racionales, de valor aparente o de sentido común. Los instrumentos derivados racionalmente se elaboran seleccionando o construyendo materiales estímulo que parecen referirse a la conducta en la que está interesado el autor de la prueba. Asimismo, un instrumento puede diseñarse de modo que sea congruente con un punto de vista teórico particular de la personalidad, y evaluar los conceptos dentro de la misma teoría, ejemplo de esto puede ser la *Escala de Preferencias Personales de Edwards*.

Ahora bien, en el *enfoque Empírico* se considera que la base de una prueba puede ser empírica; esto es, que hay una base empírica para creer que la prueba debe funcionar de la manera descrita por su autor. Cada reactivo puede seleccionarse de acuerdo con su poder para discriminar entre dos grupos de interés, como las personas con potencial para las ventas y las personas en general, o pudiera ser que se muestre una puntuación global de cierta clase para discriminar entre dos grupos. Ejemplo de este enfoque es el *Inventario Multifásico de la Personalidad de Minnesota*.

Un instrumento puede desarrollarse sobre la base de la *consistencia interna* de los reactivos. Por ejemplo, puede administrarse un número bastante grande de reactivos a un grupo de respondientes, y suponer que los reactivos cuyas respuestas estén estrechamente relacionadas o intercorrelacionadas se refieran a la

misma variable psicológica. Estas variables definidas estadísticamente se usan entonces como índices de aspectos del funcionamiento de la personalidad. Ejemplo de este enfoque es el Cuestionario de 16 Factores de Personalidad de R.B. Cattell.

Tomando en cuenta todo lo anterior a continuación se hace una descripción de los principales métodos para evaluar la Personalidad:

Las *observaciones y entrevistas* son los métodos para evaluar la personalidad que se utilizan con más frecuencia, aunque necesariamente son los más válidos, *las observaciones pueden ser controladas o no controladas, y las entrevistas estructuradas o no estructuradas*. Tanto la observación como la entrevista pueden ser más objetivas y se capacita a los observadores o a los entrevistadores, según sea el caso. También es importante recopilar datos obtenidos de observaciones y entrevistas en la forma más directa posible y prestar atención a los componentes verbales y no verbales de la comunicación interpersonal.

Algunos *procedimientos de observación especiales* incluyen el registro anecdótico, muestreo de tiempo, muestreo incidental y pruebas de situación. Un tipo especial de entrevista es la de estrés, para la cual se requiere una capacitación muy extensa a fin de que sea efectiva. Las entrevistas se utilizan para diversos propósitos como son en el reclutamiento y la selección de personal; entrevistas para diagnóstico y tratamiento psicológico, o bien para el consejo vocacional, etc.

Cuando se realiza una evaluación psicológica se emplean diversos tipos de procedimientos de *calificación* (escalas numéricas, escalas gráficas, escalas estándar, escalas con respaldo conductual y escalas de elección forzada). De todas estas, quizá la más utilizada son las escalas de calificación gráfica. Asimismo, una escala de calificación numérica de particular interés es el diferencial semántico. Las escalas elaboradas a partir de esta modalidad se utilizan para determinar los significados connotativos que los conceptos seleccionados tienen para las personas (Díaz-Guerrero, 1975).

Los errores constantes generados en las escalas de calificación como son errores de indulgencia, error de severidad, error de tendencia central y el llamado efecto halo, error lógico, error de contraste y error de proximidad, pueden reducirse mediante la capacitación extensa de los calificadores y el uso de formatos especiales para el diseño de escalas de calificación.

Una lista de verificación o cotejo de las características o problemas personales es un método popular y eficaz para registrar las observaciones por parte de uno mismo y otras personas. Existe una gran variedad de estas listas en el mercado, incluyendo instrumentos que sirven para observar y evaluar el comportamiento adaptado, la condición mental, depresión, ansiedad, problemas personales, antecedentes personales, problemas de salud, abuso sexual, y un análisis más general de la personalidad, entre las más populares están la Lista de Verificación de Problemas de Mooney y la Lista de Verificación de Adjetivos (Aiken, 1995, 1996).

Las escalas de calificación, listas de verificación e instrumentos psicométricos relacionados contribuyen en gran manera a la evaluación y comprensión de la personalidad humana, sin embargo, no tienen tanta atención e importancia como *los inventarios de personalidad y las técnicas proyectivas*. Los inventarios de personalidad se utilizan tanto en situaciones clínicas y fuera de ellas, así como también las técnicas proyectivas, sin embargo, ambas se concentran en el ajuste y psicopatología de la personalidad; se ha tratado de ampliar el diseño y la aplicación de este tipo de instrumentos a fin de incluir la evaluación e investigación de la personalidad normal, pero la importancia que daban los primeros inventarios y las primeras técnicas proyectivas a la psicopatología continua en muchos instrumentos actuales (Gregory, 1993; Aiken 1996).

Los inventarios de ajuste y temperamento, más comúnmente conocidos como inventarios de personalidad, consisten en ítemes o reactivos que se refieren a las características, pensamientos, sentimientos y comportamiento personales, en estos los sujetos marcan aquellos reactivos que juzgan más descriptivos de ellos mismos. Un instrumento así, generalmente produce calificaciones en distintas variables, así como también existen instrumentos que evalúan un sólo propósito o variable. Ejemplo de estos instrumentos son el Inventario de Depresión de Beck. Sin embargo, un inventario de personalidad típico mide una amplia diversidad de variables y por lo regular está elaborado y estandarizado con mayor detenimiento que una escala de calificación o lista de verificación.

Dentro de los primeros inventarios de personalidad creados está la Hoja de Datos Personales de Woodworth, que fue creado durante la Primera Guerra Mundial, y a partir de este se elaboraron muchos inventarios ya sea de calificación única o múltiple. Diversos tipos de inventarios están elaborados en

base a un marco de referencia personal o teórico, entre ellos esta la Escala de Preferencias Personales de Edwards, el Indicador de Tipos de Myers-Briggs. Otros instrumentos se basan en los resultados del Análisis Factorial como son el Cuestionario de 16 Factores de Personalidad, el Estudio del Temperamento de Guilford-Zimmerman y el Cuestionario de Personalidad de Eysenck (Kline, 1993; Aiken, 1996).

El inventario de personalidad más famoso, y que más investigación reporta es el Inventario Multifásico de la Personalidad de Minesota (MMPI), dicho instrumento está basado en la codificación de criterios, fue diseñado para diferenciar entre distintos grupos al analizar las diferencias en las respuestas en nueve escalas clínicas de personas normales y pacientes con diagnóstico psiquiátrico específico. Actualmente se utiliza la versión 2 del MMPI que fue reestandarizada a partir de una primera versión en 1986. El procedimiento llamado de codificación de criterios con el cual se elaboró el MMPI también se utilizó en la creación del Inventario Psicológico de California, el Inventario de Personalidad para Niños y otros instrumentos, que incluyen el Inventario Multiaxial de Teodoro Millon (MCMI) (Buros, 1992; Aiken, 1995, 1996; Groth-Marnat, 1990).

Los instrumentos de evaluación de la personalidad mejor conocidos como Técnicas Proyectivas, se crearon a partir de un término que creó Frank en 1939 (citado por Gregory, 1992), para los estímulos ambiguos en base a que los sujetos puedan "proyectar" sus necesidades y sentimientos internos. Las técnicas proyectivas están compuestas por estímulos relativamente no estructurados y/o tareas que se pide al sujeto que describa, cuente una historia, complete o responda de alguna otra manera. Los psicólogos que apoyan las técnicas proyectivas sostienen que los inventarios e instrumentos de autodescripción no muestran aspectos profundos de la personalidad ya que estos están demasiado delimitados en las respuestas. Debido a que la calificación de las técnicas proyectivas es muy subjetiva, estas enfrentan dificultades en lo referente a la confiabilidad y la validez.

Las técnicas proyectivas pueden agruparse en cinco categorías, de acuerdo al tipo de respuesta requerida: *técnicas de asociación* (palabras o manchas de tinta), *técnicas de construcción* (historias), *técnicas de completamiento* (frases o historias), *técnicas de arreglo o selección* (de dibujos u opciones verbales), y *técnicas de expresión* (dibujos, juegos). Ejemplo de técnicas de asociación están la Asociación de Palabras, Recuerdos Inmediatos, y la Prueba de Manchas de

Tinta de Rorschach. Dentro de las técnicas de construcción está el Test de Apercepción Temática, y la técnica de Casa-Árbol-Persona. Las Frases Incompletas y la Prueba de Frustración de Rosenzweig, serían buenos ejemplos de las técnicas de completamiento. En las técnicas de arreglo o selección están el Test de Szondi, y la Prueba de Arreglo de Símbolos de Kalm. Por último tenemos dentro de las técnicas de expresión el Dibujo de la Figura Humana de Machover. (Aiken, 1996; Gregory, 1992).

Todo lo anterior permite tener un bosquejo general de todo lo referente a la Personalidad, cómo se define y clasifica, y además cómo se ha tratado de medir. Sin embargo, gracias a las computadoras y la tecnología se sigue haciendo investigación amplia tratando de encontrar instrumentos y procedimientos cada vez más finos y efectivos que permitan por un lado entender la conducta del ser humano y por otro predicciones más precisas y eficaces.

Capítulo II

TESTS PSICOLÓGICOS

TESTS PSICOLOGICOS.

2.1. Historia de la medición en psicología.

El origen de los tests psicológicos surgió debido al interés por conocer las diferencias humanas, por un lado, y para poder tomar decisiones a partir de los datos obtenidos. (Nunnally, 1970; Pichot, 1976; Anastasi, 1988; Cerdá, 1984). Sin embargo, se tienen datos sobre los primeros intentos de estudiar a los seres humanos en sus diferencias y semejanzas desde la época de los mandarines chinos que implementaron un servicio civil para determinar si sus funcionarios eran aptos para desempeñar sus deberes gubernamentales; por otro lado, tanto Platón como Aristóteles ya hablaban hace 2 500 años sobre las diferencias individuales (Aiken, 1995; Gregory, 1993).

En el intento por conocer las diferencias humanas se realizaron diversas investigaciones que van desde el estudio de las sensaciones del ser humano hasta el nivel intelectual y las emociones de cada persona.

Estas investigaciones tuvieron gran importancia dentro de la psicología ya que llevaron a la creación de las pruebas o tests psicológicos, que hasta la fecha permiten conocer diversas características de los seres humanos.

Entre los primeros estudios de las diferencias humanas se encuentra el trabajo del español J. Huarte de San Juan quien, en el siglo XVI, daba normas para descubrir habilidades e ingenios y establecer diferencias mediante su *Examen de ingenios*.

En el siglo XIX en Francia, ante el trato inadecuado que sufrían los enfermos mentales surge el interés por estudiarlos y hacer clasificaciones que permitieran poder darles un tratamiento adecuado. En el año de 1838, el psiquiatra Esquirol publicó: *Des maladies mentales considerees sous les rapports medical hygienique et medico-legal*, donde realizó la primera clasificación de los enfermos mentales, la cual se basó en el uso del lenguaje que podía tener un sujeto, ya que consideraba que éste era un criterio adecuado para determinar el nivel intelectual (Cerdá, 1984).

Es importante mencionar que el trabajo que realizó Esquirol sirvió como referencia para la construcción de las primeras escalas de inteligencia y desarrollo, como las de Binet-Simon y Gessell.

Durante ese mismo siglo Ernest Weber, en el año de 1840, publica los resultados de sus trabajos sobre relaciones entre estímulos y respuestas sensoriales donde postula diversas leyes, entre las cuales destacan las del "*umbral absoluto y umbral diferencial*" (Cerdá, 1984).

Fechner se interesó y continuó los trabajos de Weber; sus investigaciones se enfocaron a medir las sensaciones por medio de los juicios humanos, de esta manera llegó a la conclusión de que la "*respuesta sensorial es proporcional al logaritmo del estímulo*" (Cerdá, 1984).

Wundt, en el año de 1879, crea el primer laboratorio de psicología experimental, que tenía como finalidad formular leyes sobre la conducta humana; en este laboratorio se trabajó con diversos sujetos observando que no todos reaccionaban de la misma manera, por lo que no se podían realizar generalizaciones, por lo tanto, fue necesario que todos trabajaran bajo las mismas circunstancias y se estudiaran entonces las diferencias individuales. Los experimentos que ahí se realizaban tenían influencia médica y fisiológica, ya que se hacían estudios sobre las reacciones de los sentidos ante diversos estímulos.

Posteriormente Francis Galton, en Londres, se dedicó a hacer estudios sobre la herencia humana realizando mediciones en personas que tenían o no parentesco. Galton, influenciado por Locke, creó diversos tests para medir funciones sensoriomotoras ya que consideraba que el conocimiento entra por los sentidos, y por lo que la medición de estas funciones le permitiría conocer la inteligencia de cada ser humano. En el año de 1884 estableció un laboratorio antropométrico, en el Museo de South Kensington, donde a los visitantes se les aplicaba una serie de pruebas que les permitía conocer su altura, peso, capacidad auditiva y otros atributos. Se emplearon por primera vez las escalas de estimación y los cuestionarios.

Una de las aportaciones de Galton fue el uso de procedimientos estadísticos en los datos que reportaban el uso de los tests; dichos procedimientos los desarrolló con su discípulo K. Pearson; quien estableció la fórmula del coeficiente de correlación, la correlación parcial, correlación múltiple y el análisis factorial.

En 1904 Spearman establece el uso del análisis factorial al utilizar el coeficiente de correlación de Pearson en los resultados de diferentes tests.

James McKeen Cattell, un psicólogo norteamericano que fue colaborador de Galton y discípulo de Wundt, siguió los estudios de las diferencias individuales y el uso de los tests en los Estados Unidos. En el año de 1890 publicó un artículo llamado *Mental Test and Measurements*, en la revista *Mind* (Cerdá, 1984), donde utilizó por primera vez el término **test mental** para designar a una serie de pruebas o instrumentos psicológicos que se empleaban para conocer las diferencias individuales y la inteligencia. Cattell, consideraba al igual que Galton, que la inteligencia se podía medir a través de la discriminación sensorial y el tiempo de reacción.

Kraepelin, quien es considerado el padre de la Psiquiatría, en el año de 1895 crea tests para explorar funciones más complejas como la elaboración de conceptos, enjuiciamiento práctico, fatiga mental, etcétera.

En 1897 Ebbinghaus elabora instrumentos para medir **memoria** y cálculo aritmético.

Los italianos Guicciardi y Ferrari, continuaron el trabajo de Kraepelin y diseñaron pruebas destinadas a la exploración de los enfermos mentales.

A finales del siglo XIX A. Binet y Herni realizan un trabajo llamado *La psychologie individuelle* donde se oponen totalmente a la teoría de Locke y sus seguidores, que sostenía que la inteligencia se podía medir a través de los sentidos, argumentando que ésta no tomaba en cuenta las diferencias individuales. Propusieron entonces diversos tests que incluían la medición de la memoria, la comprensión, la imaginación y la atención.

En 1904 se formó una comisión dentro de la Institución Pública francesa para estudiar las formas de educación de los niños subnormales y el diagnóstico de la enfermedad mental.

Binet y Simon publican en 1905 en *L'Année Psychologique* un artículo llamado "Nuevos métodos para el diagnóstico del nivel intelectual de los anormales" (Pichot, 1976), donde se incluía el primer **test mental práctico**. Esta escala contaba con problemas ordenados en grado de dificultad y de acuerdo al ambiente cultural. Sus subtests medían el juicio, la comprensión y el razonamiento, dentro de sus reactivos se incluía el nombrar objetos, completar

frases o comprender preguntas. Este test se presentó como un instrumento provisional y no contaba con ningún método para su calificación.

En 1908 apareció la primera revisión de la escala Binet-Simon: se eliminaron los tests que no eran satisfactorios de la versión anterior y se agregaron nuevos tests. Además se realizó una clasificación de los tests de acuerdo a diversos niveles de edad. En esta escala se utilizó por primera vez el término "*Edad mental*", que correspondía a los niños cuya ejecución igualaba la edad cronológica (Anastasi, 1988).

La tercera revisión de esta escala se realizó en 1911: en ésta se le adicionaron varios niveles de edad, contemplando su uso en adultos.

La escala Binet-Simon despertó un gran interés en diversas partes del mundo, se realizaron diversas traducciones y adaptaciones, una de las más importantes fue la que llevó a cabo L.M. Terman en la Universidad de Stanford en 1916, conocida como *Escala Stanford-Binet*, en la cual Stern sugirió el término "*Coficiente Intelectual*", que consideraba la relación entre la edad mental y la edad cronológica (Anastasi, 1988).

Las diversas formas de la escala de Binet, requerían en su aplicación un examinador que interactuará con cada sujeto.

En 1917 cuando el gobierno de los Estados Unidos, decide participar en la Primera Guerra Mundial, se enfrenta ante el problema de seleccionar rápidamente a un millón y medio de personas para el ejército y elegir su cargo por lo que el *Ministerio del Ejército* forma un comité dirigido por R. K. Yarkes para investigar acerca de todos los procedimientos disponibles para elaborar dicha selección.

Durante esa época Arthur S. Otis había creado un test colectivo de inteligencia; formado con preguntas de información general, razonamiento y aritmética, que presentó ante el comité, en donde se revisó y adaptó para poderse emplear en la selección de los reclutas. A este test se le conoció como "*Test Alfa*" ("*Army Alpha Test*") (Cerde, 1984). Se construyó además un test complementario para personas analfabetas y con poco dominio del idioma inglés, que incluía rompecabezas geométricos y análisis de ilustraciones, el *Army Beta*. Estos tests fueron cedidos para posterior uso civil.

Durante la Primera Guerra Mundial, se desarrollaron también tests de personalidad, dentro de los cuales destaca el que elaboró Woodworth para los

reclutas y que servía para detectar casos que podrían requerir un mayor estudio clínico. Este test estaba formado por preguntas que se emplearían comúnmente en una entrevista referentes a la adaptación del sujeto en su hogar, en la escuela y con sus amigos.

En Suiza, también durante la Primera Guerra Mundial, el psiquiatra Hermann Rorschach implementaba una nueva forma para evaluar la personalidad. Elaboró una prueba en donde al sujeto se le presentaban 10 láminas con manchas de tinta y se le pedía que las describiera. Esta técnica tenía influencia psicoanalítica y daba relevancia al aspecto simbólico.

En 1939 aparece la primera Escala Wechsler-Bellevue, que sirve para medir inteligencia, siendo ésta la primera escala para adultos en la que se llevó a cabo una estandarización en función de su edad cronológica.

Anastasi (1988) menciona que el desarrollo de los tests psicológicos, en un principio se orientó principalmente a medir diversas funciones y a partir de éstas determinar el nivel de inteligencia, estos instrumentos medían en su mayoría aptitud verbal y en menor grado aptitudes para manejar relaciones numéricas, abstractas y simbólicas. Posteriormente se descubrió que estas medidas proporcionaban poca información de cada función, por lo que fue necesario crear tests que midieran funciones específicas.

Desde antes de la Primera Guerra Mundial los psicólogos comenzaron a crear tests que midieran aptitudes especiales que complementarían a los tests de inteligencia, se crearon instrumentos para conocer las aptitudes profesionales, para el reclutamiento y la selección de personal, entre otras.

Al mismo tiempo en que se descubría la necesidad de elaborar tests para medir las diversas aptitudes de los individuos, se estudiaba la organización de los rasgos, por lo que se facilitó la elaboración de éstos. En esta área, Spearman realizó una de las mayores aportaciones con sus investigaciones estadísticas en estudios de intercorrelaciones de la naturaleza de la inteligencia y las puntuaciones en otras pruebas; y sus estudios sobre el análisis factorial, en donde se descubrió que en las puntuaciones de una prueba, están presentes diversos factores o rasgos independientes.

Los estudios realizados con el análisis factorial permitieron el desarrollo de baterías de aptitud múltiple, que tenían como finalidad proporcionar una medida de la posición del sujeto en cierto número de rasgos.

En 1941, se publican los Tests de Chicago que medían aptitudes mentales primarias (PMA), siendo este el primer esfuerzo por construir una batería.

Fue después de la Segunda Guerra Mundial, en 1945, cuando los psicólogos militares realizaron investigaciones sobre los tests basándose en el análisis factorial dirigidos a la elaboración de instrumentos para medir aptitudes diferenciales. Durante este período se crearon tests especiales para pilotos, bombarderos, operadores de radio dentro de la Army Air Forces 1947-1948.

Por otro lado en el ámbito escolar se experimentaban nuevas técnicas, dentro de las cuales se encontraba la aplicación de pruebas tipificadas de rendimiento, siendo éstas las primeras que consideraban importante que todos los sujetos se sometieran a las mismas circunstancias. En estas pruebas se trataba de abarcar contenidos más amplios y eliminar la posibilidad de favoritismo por parte del examinador durante el proceso de evaluación (Mann, citado en Anastasi, 1988).

Durante esta misma época se establecieron programas de tests a escala estatal, regional y nacional, de los cuales el más conocido fue el que realizó el *College Entrance Examination Board* (CEEB). Posteriormente los servicios del CEEB se unieron al *Carnegie Corporation* y los del *American Council on Education* formando el *Educational Testing Service*, quienes se encargaron de elaborar pruebas para universidades, escuelas profesionales, y organismos gubernamentales (Anastasi, 1988).

En 1959 el *American College Testing Program* elaboraba instrumentos para los aspirantes de ingreso en los centros no incluidos en el CEEB.

Dentro del estudio de la personalidad, se crearon tests para evaluar los intereses, actitudes y opiniones, dentro de los primeros que se elaboraron se encuentra el que fue creado en el seminario para graduados celebrado en el Instituto de Tecnología Carnegie 1919-1920 (Fryer, 1931, citado en Anastasi, 1988), esta prueba se denominaba Strong Vocational Interest Blank (SVIB). El grupo del Carnegie creó diversos inventarios de intereses los cuales presentaban preferencias o aversiones de los sujetos para una variedad de actividades específicas, objetos o personas que se encuentran comúnmente en la vida diaria (Anastasi, 1988). Posteriormente se crearon otros inventarios, dentro de los cuales destaca la elaboración del MMPI (1939) (Inventario Multifásico de la Personalidad Minnessota) y el CPI (*California Psychological Inventory*).

Anastasi (1988) señala que para complementar el SVIB, se elaboró un inventario similar de intereses para diversas ocupaciones especializadas y semiespecializadas. Este inventario se llamó *Minnesota Vocational Interest Inventory*, formado por 21 escalas para ocupaciones tales como panaderos, conductores, reparadores de televisión, entre otros.

Otro de los inventarios más conocidos para evaluar intereses son las diversas formas y versiones del Kuder, dentro de las cuales se encuentra el *Kuder Preference*, el *Kuder Occupational Interest Survey*, y el *Kuder General Interest Survey*.

En lo referente a la historia de la medición en psicología es importante señalar que aquí solamente se han citado algunas de los tests que marcaron un paso importante en la historia, y que el desarrollo de los tests continua hasta la fecha.

2.2. Definición de los tests psicológicos.

La conceptualización y definición de las pruebas o tests psicológicos varía de acuerdo a cada autor, a continuación se presentará una definición breve de los autores más relevantes:

Tyler, L. (1972) señala que: "*Una prueba puede definirse como una situación estándar diseñada para tomar una muestra del comportamiento de un individuo*".

Pichot, P. (1976) da la siguiente definición: "*Un test es una situación experimental estandarizada, sirviendo de estímulo a un comportamiento.*"

Anastasi (1988) considera que un test tiene como finalidad medir las diferencias que existen entre diferentes individuos o las reacciones del mismo individuo en diferentes ocasiones. Por lo tanto, Anastasi define: "*Un test psicológico constituye una medida objetiva y tipificada de una muestra de conducta*".

Brown (1980) define a una prueba como el: "*Procedimiento sistemático para medir una muestra de conducta*", "*El procedimiento sistemático se refiere a que una prueba se construye, se administra y se aplica, según las reglas preestablecidas*".

"Test" es una palabra inglesa que significa "prueba" y que se deriva del latín "testis". Esta misma raíz figura en palabras españolas tales como testimonio o testigo. La palabra "test" se utiliza sin traducir en todos los países del mundo y sirve para designar una modalidad de exploraciones muy extendida hoy en día en diversos campos científicos y técnicos entre ellos el psicológico (Cerdá, 1978).

Cerdá (1978) menciona la definición que fue aceptada en la Antigua Asociación Internacional de la Psicotécnica de H. Piéron que dice: "*Test es una prueba definida, fijando una tarea a realizar, idéntica para todos los sujetos examinados. Debe disponer de una técnica precisa, que permita discriminar entre soluciones satisfactorias y erróneas o bien de un sistema numérico que permita puntuar el resultado. La tarea puede consistir en poner de manifiesto conocimientos adquiridos (tests pedagógicos) o bien funciones sensorio-motoras o mentales (tests psicológicos).*" Cerdá señala que esta definición sólo incluye tests de inteligencia, aptitudes y conocimientos, y no toma en cuenta los instrumentos de personalidad.

Considerando la definición de los diversos autores, se puede concluir que una prueba o test psicológico es una situación estandarizada, que permite tener una medida objetiva de la conducta de un individuo.

2. 3. Los reactivos o ítems que conforman un test.

En el proceso de elaboración de una prueba o test, se deben de tomar en cuenta como elemento básico de estos los llamados reactivos o ítems que lo conforman. Para ello autores como Aiken (1995), Kline (1986), Reidl (1990), hacen las siguientes sugerencias para la elaboración de reactivos o ítems:

- Se deberán usar palabras sencillas, que sean conocidas por todos los respondientes potenciales.
- Hacer la pregunta lo más concisa posible, esta no deberá tener más de 20 palabras.
- Formular la pregunta de manera que produzca únicamente la información deseada, evitando reactivos que no estén relacionados con la variable.
- Evitar preguntas ambiguas o que puedan ser interpretadas de diferente manera.
- Evitar preguntas que sugieran la respuesta.

- Cada reactivo o pregunta debe tener sólo una idea, oraciones simples.
- La sintaxis debe ser sencilla y clara: sujeto, verbo y complemento.
- Se debe tener cuidado de no emplear frases que pueden afectar el prestigio del respondiente.
- Decidir si se van a personalizar algunos de los reactivos o ítemes, o si se redactarán en tercera persona o en forma impersonal.
- Permitir que se den todas las posibles respuestas, principalmente en los reactivos de opción múltiple.
- Las alternativas de los reactivos de opción múltiple deben ser realistas.
- Emplear con cuidado o casi nunca palabras universales como: "*siempre*", "*ninguno*", "*nunca*"; (solo para definir los extremos de continuos) o aquellas como: "*solo*", "*apenas*", "*casi*", "*a veces*", porque introducen ambigüedad ya que tienen diferentes significados para las personas.
- Se debe evitar el usar dobles negativos, ya que también producen ambigüedad.
- Tratar de mantener al mínimo la cantidad de escritura por parte del que responde, pidiéndole tan solo que escriba SI o NO; escriba un número o una letra; ponga una cruz (X); subraye la respuesta que desee, circule la respuesta escogida, etc.
- Evitar en lo posible términos dentro del reactivo que reflejen sentimientos.
- En los reactivos de opción múltiple coloque las opciones en formato (párrafo) de grupo en lugar de una tras otra; utilice números para designar los reactivos y letras para las opciones.

2.4. Tipos de reactivos o ítemes.

Los reactivos que conforman un instrumento representan procedimientos para obtener información sobre los individuos, pero la cantidad y clase de información varía con la naturaleza de las tareas que implican los distintos tipos de reactivos, Aiken (1995). Si por ejemplo le pedimos a un sujeto que compare la batalla de la Noche Triste, con la batalla de la toma de Tenochtitlán esto exige una clase de respuesta diferente, que aquella que se obtiene si le pedimos al sujeto

que indique una serie de eventos que ocurrieron en cada batalla. El primer reactivo requiere de capacidades complejas de integración y organización, en el segundo sólo se necesita memoria cognoscitiva. Ahora bien, existen reactivos de ensayo o abiertos y reactivos objetivos que pueden ser de respuesta corta o completamiento, de falso y verdadero, de apareamiento y de opción múltiple. A continuación se describen estos diferentes tipos de reactivos:

Reactivos de Ensayo o Abiertos: En este tipo de reactivo se le pide al sujeto que elabore su respuesta de manera libre. Una de la ventajas de este tipo de reactivos es que pueden medir la capacidad para organizar, relacionar y comunicar comportamientos que no evalúan con facilidad los reactivos objetivos. Ahora bien, una desventaja de este tipo de reactivos es que la cantidad de preguntas que pueden responderse en un tiempo de clase regular tal vez no proporcione un muestreo adecuado del conocimiento que sobre la materia tienen las personas. No deben utilizarse este tipo de reactivos cuando es posible realizar la misma evaluación con reactivos objetivos. Al redactar estos reactivos es necesario que la persona realice preguntas objetivas. Esto se puede lograr al:

- 1) definir la tarea y redactar los reactivos de forma clara
- 2) utilizar una cantidad reducida de reactivos que deberán responder todas las personas;
- 3) estructurar los reactivos de tal manera que los expertos en la materia estén de acuerdo en que una respuesta es mejor que otra de forma demostrable;
- 4) pedir a las personas que respondan en una hoja de papel separada. Ejemplo:

Instrucciones: Escriba una respuesta de media página para cada reactivo.

- 1. Compare las ventajas y desventajas entre los reactivos abiertos y los objetivos.**
- 2. Explique las razones de por qué es necesario realizar análisis de reactivos cuando se crea un instrumento.**

Reactivos de Respuesta Corta o Completamiento: Es una tarea en la que se le pide al sujeto complete o llene uno o mas espacios en blanco de una afirmación incompleta con las palabras o frases correctas o que den una respuesta breve a una pregunta. Son reactivos que se elaboran con mucha facilidad, ya que se requiere que las personas den la respuesta correcta en lugar de reconocerla. Ejemplo:

Instrucciones: Escriba la (s) palabra (s) apropiada (s) en cada espacio.

1. El creador del 16 PF fue _____.

2. Confiabilidad se refiere a _____.

Reactivos de Falso y Verdadero: En este tipo de reactivos se pide al sujeto que únicamente indique la respuesta, ya sea verdadera o falsa. Pueden redactarse y leerse con rapidez, y permiten un muestreo extenso del contenido. Una desventaja importante de este tipo de reactivos es que se ocupan de información trivial, o se elaboran con afirmaciones que se tomaron al pie de la letra de los libros de texto. Se dice que alientan el aprendizaje de memoria. Ejemplo:

Instrucciones: marque con una cruz la letra V si la afirmación es verdadera; marque con una cruz la letra F si la afirmación es falsa.

V F 1. Un instrumento objetivo tiene confiabilidad y validez.

V F 2. El 16 PF es un test de inteligencia.

Reactivos de Aparejamiento: En este tipo de reactivos se compara un grupo de opciones de estímulo o pregunta con un grupo de opciones de respuesta. Por lo general la comparación es de uno a uno, pero también puede ser de uno a varios. Son fáciles de elaborar y cubren el material con mas eficacia que otros tipos de reactivos. Un tipo especial de este tipo de reactivos sería el de *reordenamiento o jerarquización*, en el cual se pide a las personas que clasifiquen un grupo de opciones en una cantidad fija de categorías determinadas previamente. Ejemplo:

___ 1. Test de Inteligencia Grupal.	A. Rorschach
___ 2. Test de Inteligencia Individual.	B. Stanford
___ 3. Inventario de Intereses.	C. MMPI
___ 4. Inventario de Personalidad.	D. Kuder
___ 5. Test de Manchas de Tinta.	E. Raven

Reactivos de Opción Múltiple: La tarea consiste en elegir de una serie de respuestas la correcta para una pregunta dada. Son los más versátiles de todos los reactivos para instrumentos de tipo objetivo, ya que pueden utilizarse para medir los objetivos de aprendizaje complejos y sencillos en todos los niveles y en cualquier materia. Responder a un reactivo de opción múltiple correctamente elaborado requiere de gran capacidad para la discriminación y no sólo para reconocer o recordar la respuesta adecuada. Ejemplo:

Instrucciones: Marque la opción correcta en su hoja de respuestas de acuerdo a la pregunta que vaya contestando.

1. Es el inventario de personalidad más utilizado en la historia de la psicología.

- A. 16 PF
- B. TJTA
- C. MMPI
- D. CPI

2. El Test de Raven es una prueba de:

- A. Personalidad
- B. Intereses
- C. Valores
- D. Inteligencia

En el caso de los instrumentos objetivos de personalidad, existen también los siguientes tipos de reactivos: *reactivos abiertos*, *cerrados*, *en abanico de respuesta*, *de estimación*, y *de relleno* (Hernández, M.S., León, B.L, Martínez, F.J., 1985).

Los *reactivos abiertos* son aquellos que dejan un margen de libertad total para responder al evaluado; deben de estar redactados con gran claridad y tener una sola interpretación posible para que las respuestas que se obtengan sirvan efectivamente a la investigación (es decir, que respondan a aquello que interesa conocer). La información que proporcionan este tipo de reactivos es muy rica, pero plantean dificultades para el análisis de datos, debido a que, para su codificación y tabulación, es preciso realizar un análisis de contenido a través del cual, se puedan determinar las categorías que se han manifestado. Ejemplo:

¿Cómo cree usted que debe reaccionar un padre cuando descubre que su hijo le ha mentado?

Los *reactivos cerrados* son aquellos donde la opción de respuesta se da entre la afirmación y la negación; estos son de redacción simple y de gran utilidad por la rapidez con que se pueden procesar, pero no brindan, desde luego, una información muy amplia. Ejemplo:

¿Padece usted de insomnio?

SI () NO ()

Existen dos derivaciones del tipo de *reactivo cerrado* que son: los *reactivos en abanico de respuestas* y los *reactivos de estimación*.

Los *reactivos en abanico de respuestas* son aquellos donde el examinado debe ubicar su contestación dentro de una serie de alternativas propuestas por el investigador. Ejemplo:

¿Cuál es su estado civil?

- () Casado
- () Soltero
- () Viudo
- () Divorciado

Si bien este tipo de reactivos es de redacción más o menos sencilla, plantean el problema de que es posible que el investigador, al prever las alternativas de respuesta, este proyectando sus propias opiniones y prejuicios. En algunos casos esto no se plantea pero cuando se estima que puede llegar a ocurrir el investigador puede agregar las alternativas que él propone: la de "*otros tipos de respuestas*" y limitar en algo los efectos de su influencia. Ejemplo:

¿Porque eligió la carrera de maestro?

- () Porque le gusta tratar con niños
- () Porque le gusta enseñar
- () Porque era la única que se impartía en su pueblo
- () Porque es una manera de continuar estudiando siempre
- () Otros

Los *reactivos de estimación* son aquellos en los cuales las alternativas de respuesta se dan en forma de escala. Por ejemplo, cuando las respuestas posibles son : muy favorable, favorable, neutral o desinteresado en la cuestión, desfavorable y muy desfavorable. Los problemas que plantean estos reactivos son similares a los de las escalas de opinión y actitud, y se centran en torno a la cuestión de si lo que una persona valora, por ejemplo, como muy favorable, es lo mismo que lo que valora otra, o si existen diferencias entre ambas. Esta cuestión continua creando polémicas pero no deja por ello de ofrecer una perspectiva que, *aceptada como tendencia*, pueda resultar de gran utilidad. Ejemplo:

¿Cuál es su opinión sobre la terapia centrada en el cliente?

- () Cree que es el mejor método psicoterapéutico
- () Cree que es un buen método psicoterapéutico
- () Desconoce este método o le es indiferente
- () Cree que es un mal método psicoterapéutico
- () Cree que es el peor método psicoterapéutico

Los *reactivos test* son los que se formulan varias veces en torno a una misma cuestión, aunque pueden redactarse de manera diferente, y su objetivo es determinar si el examinado ha respondido con veracidad. Resultan particularmente útiles cuando el investigador estima que existen inhibiciones en los examinados tendientes a desviar las respuestas. Ejemplo:

¿Les pega a sus hijos cuando cometen travesuras?

¿Qué opina de los castigos corporales?

¿Cómo reaccionaría si sus hijos se pusieran a dibujar en las paredes de su casa?

Una alternativa a los *reactivos test* son las *preguntas cargadas*, llamadas así porque contienen un marcado prejuicio que precipita y define la posición del examinado (en este caso no es necesaria la repetición).

Finalmente, pueden existir dentro de un instrumento algunos *reactivos de relleno*, cuya función es precisamente la de servir de aislante entre reactivos que podrían sugerir contagio o que podrían despertar desconfianza en el examinado, o bien, preparar para responder sobre una cuestión verdaderamente difícil.

2. 5. Características de los tests psicológicos.

Toda prueba, test o instrumento de medición psicológica, debe cumplir con características específicas, para que los datos que ésta proporcione sean objetivos, de utilidad y confiables. De ahí que un test debe *ser sensible, objetivo, confiable y válido*. *Sensible* ya que los reactivos que componen el instrumento puedan diferenciar entre aquellos sujetos que tiene el rasgo de aquellos que no. *Objetivo* se refiere a que dos aplicadores o más obtengan el mismo puntaje al calificar el instrumento. *Confiable*, que sea lo más preciso en el rasgo que está midiendo. *Válido*, que mida lo que pretende medir.

2.5.1. Grado de dificultad y poder discriminativo de los reactivos o ítems.

Una de estas características es el *grado de dificultad* (Anastasi, 1988) o la *sensibilidad de discriminación* (Pichot, 1976) que deben tener los *reactivos o ítems* que forman una prueba o test psicológico. Esta característica permite conocer las diferencias individuales; y se logra mediante un procedimiento que se lleva a cabo en el momento de construir el test. El procedimiento consiste en agregar *reactivos o ítems* que puedan ser contestados por todos los sujetos y *reactivos o ítems* que sólo puedan ser contestados por aquellos más aptos. Generalmente en las pruebas psicológicas que evalúan el rendimiento académico, inteligencia y aptitudes, los reactivos son ordenados en grado de dificultad, colocando primero los reactivos más fáciles y posteriormente los de mayor dificultad, esto se hace con el fin de que los sujetos sientan confianza al empezar la prueba y permitir que los sujetos con menor capacidad, puedan dedicar mayor tiempo a aquellos reactivos que pueden resolver correctamente.

El poder discriminativo describe el grado en que la probabilidad de la respuesta alfa (aprobar un reactivo, calificando como correcto contra incorrecto, responder un reactivo en dirección clave) se correlaciona con el atributo. Un reactivo que no discrimina debe ser eliminado del instrumento. El índice clásico más común de la discriminación es la correlación con personas entre la respuesta alfa y el puntaje total de la prueba, la correlación reactivo total. (Nunnally, 1995).

2.5.2. Estandarización

La *tipificación o estandarización* es otra de las características que debe incluir una prueba o test psicométrico, ésta consiste en utilizar los mismos procedimientos para aplicar y calificar los instrumentos en todos los sujetos. La

tipificación abarca pues, materiales, límites de tiempo, instrucciones, demostraciones previas, formas de resolver dudas a los sujetos durante la aplicación. El hecho de tener bajo las mismas condiciones a los sujetos permitirá, posteriormente generar normas que permitan hacer comparaciones entre éstos, así como el tratar de minimizar al máximo errores durante la ejecución.

Dentro de las características más importantes que debe cubrir una prueba o test psicométrico es el que éste tenga un grado aceptable de confiabilidad y validez. A continuación se presentará una breve definición de ambos términos y posteriormente se revisará detalladamente cada uno de éstos.

La *confiabilidad* de una prueba o test "se refiere a la consistencia o estabilidad de las puntuaciones obtenidas de un mismo sujeto en diferentes ocasiones o con conjuntos distintos de elementos equivalentes" (Anastasi, 1988).

La *validez*: "Es el grado en que un test psicológico mide lo que realmente pretende medir" (Anastasi, 1988), es decir, los elementos que conformen la prueba o test psicométrico, deben ser representativos de la conducta que se desea medir.

2.5.3. Confiabilidad

Al realizar diversas mediciones sobre un mismo objeto, es muy probable que se cometan errores entre una y otra medición, por lo que éstas pueden no ser exactas o verse alteradas. Al medir atributos psicológicos, nos encontramos con que muchos de estos no son estables, y en ocasiones aunque se utilice el mismo instrumento en distintas ocasiones las puntuaciones obtenidas de un mismo sujeto pueden variar. Por tal motivo, es importante determinar si las variaciones en las puntuaciones de los sujetos se deben a un cambio real en el atributo medido o a un error cometido durante el proceso de medición.

Para tal efecto, la teoría psicométrica contempla el estudio del *error de medición* al calcular la confiabilidad de una prueba o test. Se considera *error de medición*, a todas aquellas condiciones ajenas o imprevistas que se presentan durante el proceso de medición y que no fueron contempladas cuando se diseñó la prueba o test psicométrico.

Nunnally (1995) menciona que *el error de medición* es una combinación de procesos sistemáticos y procesos aleatorios. Los procesos sistemáticos son las fallas en el procedimiento, que incluyen desde la elaboración, aplicación y

calificación de la prueba, hasta la captura y el análisis de los datos. Los procesos aleatorios son los que quedan fuera de control del proceso de la medición, como los cambios propios del elemento medido.

Si el *error de medición* es sistemático y altera todas las mediciones por igual, es considerado como un *error constante*, que no tendrá mayor efecto en la medición puesto que no alterará las comparaciones de los grupos. Pero cuando el error sólo afecta algunas de las mediciones se considera que hay un sesgo en la medición.

El *error aleatorio* tiene mayor importancia en las mediciones psicológicas puesto que los puntajes obtenidos se ven afectados por diversas situaciones tales como la elección adecuada de los reactivos de acuerdo a la muestra de sujetos, es decir, determinada prueba puede arrojar puntuaciones elevadas en un grupo y bajas en otro, probablemente porque uno de estos tuvo mayor acceso a información que le fue útil al contestar la prueba y el otro grupo no; otra situación que puede variar las puntuaciones de un mismo sujeto, es el empleo del azar al contestar los reactivos de la prueba o que no tenga la misma disposición en las diversas mediciones. Este tipo de error es casi imposible de eliminar, por lo que se debe de tratar de reducir al mínimo siempre que sea posible.

Debido a la cantidad de errores que se pueden cometer en las mediciones es importante conocer la confiabilidad de una prueba. Anastasi (1988) define la *confiabilidad* de una prueba o test como "*la consistencia de las puntuaciones obtenidas por los mismos individuos cuando son examinados con el mismo test en diferentes ocasiones, con conjuntos distintos de elementos equivalentes o bajo otras condiciones variables del examen*".

Para obtener la confiabilidad de una prueba o test psicométrico se obtiene un *coeficiente de correlación*, que indica el grado de concordancia que hay entre dos conjuntos de calificaciones. Para estimar el *coeficiente de correlación* se emplean algunos métodos estadísticos como la fórmula de correlación de Pearson, o Spearman entre otros.

Los valores que puede asumir una correlación van desde -1 hasta +1, cuando el valor obtenido de la correlación es positivo se dice que se ha encontrado una correlación positiva, y cuando el valor es negativo la correlación se denomina negativa. Cuando el valor del coeficiente se acerca más a +1 se considera que la correlación es significativa. Es importante señalar que los coeficientes de correlación que se encuentran en una prueba psicológica son

generalmente positivos. El coeficiente de correlación toma en cuenta la posición del sujeto en el grupo así como la distancia que hay en el puntaje de cada sujeto respecto al promedio de calificaciones obtenidas. Es importante señalar que en las mediciones de atributos psicológicos es imposible llegar a obtener un grado de correlación igual a 1. Este coeficiente también es denominado por algunos autores como *coeficiente de confiabilidad*.

Debido a los diversos factores que pueden alterar la confiabilidad, se pueden considerar diferentes tipos de confiabilidad, algunos de éstos contemplan el tiempo que transcurre entre una y otra aplicación como es el caso del *coeficiente de estabilidad*, otros consideran la adecuada elección de los reactivos que forman la prueba y proporcionan el *coeficiente de consistencia interna* y/o el *coeficiente de equivalencia*. Cada uno de éstos presenta una varianza de error distinta, en seguida se describirá brevemente en que consiste cada uno de estos métodos así como el tipo de coeficiente de correlación que aporta y el tipo de varianza de error que contempla.

Métodos para obtener la Confiabilidad

2.5.3.1. Método de confiabilidad test-retest.

Este método consiste en aplicar una prueba y después de un intervalo de tiempo, volver a aplicar la misma. Posteriormente se calcula el coeficiente de confiabilidad con las calificaciones obtenidas en la primera y segunda aplicación de cada sujeto. El coeficiente de confiabilidad que se obtiene con este método se denomina *coeficiente de estabilidad temporal*. La varianza de error se debe a las variación que se puede producir en el transcurso de tiempo de una aplicación a otra, o a las condiciones de aplicación en que se administró la prueba. Es importante señalar que este tipo de confiabilidad se podría ver alterado con rasgos variables de la personalidad, como el estado de ánimo.

2.5.3.2. Método de confiabilidad con formas equivalentes o paralelas.

Consiste en elaborar dos pruebas o instrumentos parecidos que midan el mismo atributo; la construcción de pruebas paralelas consiste en crear dos formas de la prueba que contengan el mismo número de reactivos y el mismo grado de dificultad. Ambas pruebas deben tener formato, instrucciones, y límite de tiempo en la administración iguales, para posteriormente ser comparadas. Una vez construidas ambas, se aplica una forma y con un período mínimo de tiempo se aplica la segunda forma al mismo grupo de sujetos. Para estimar la confiabilidad

se correlacionan las calificaciones obtenidas en las dos formas y el resultado de ésta se denomina *coeficiente de equivalencia*. La varianza de error muestra las variaciones en la actuación de los sujetos en ambas formas. Este método permite además conocer la magnitud del error de medida debido al muestreo de contenido.

2.5.3.3. Método de confiabilidad de división por mitades.

En este método se aplica sólo una forma de la prueba o test. Después se divide el instrumento en dos mitades equivalentes que se califican por separado para obtener dos calificaciones distintas para cada sujeto. Una de las formas más comunes para dividir la prueba es separar los reactivos pares e impares y obtener una calificación sumando todos los reactivos pares y otra con los reactivos impares, en esta división se debe tener en cuenta que los reactivos pares e impares sean similares en el grado de dificultad. Posteriormente, para obtener la confiabilidad de la prueba estas dos puntuaciones se correlacionan y se obtiene un *coeficiente de consistencia interna*, se le llama así puesto que sólo emplea una forma de la prueba. Este método proporciona sólo la confiabilidad de la mitad de la prueba, pero entre mayor sea el número de reactivos o ítems que contiene el instrumento mayor será la confiabilidad.

2.5.3.4. Método de confiabilidad de Kuder Richardson.

Este método consiste en evaluar la consistencia interna de una prueba, para lo cual se requiere de una sola aplicación y la confiabilidad se estima a partir de las respuestas de los sujetos a todos los reactivos de la prueba, la forma más común de evaluar este tipo de consistencia es utilizando la fórmula Kuder-Richardson que hace un examen de la ejecución de cada elemento. Las fuentes de error que influyen en este método son el muestreo de contenido y la heterogeneidad de lo que se pretenda medir.

2.5.4. Validez.

En los tests psicológicos no sólo es importante que haya confiabilidad en las mediciones, puesto que esta sólo indica la consistencia en las puntuaciones, un problema más importante que enfrenta una prueba psicológica es comprobar empíricamente que ésta mide el atributo o rasgo para el que fue construida.

Magnusson (1973) menciona que "*la validez de un método es la exactitud con que pueden hacerse medidas significativas y adecuadas con él, en el sentido que midan realmente los rasgos que se pretenden medir*". Anastasi (1988) señala

que para poder determinar la validez es necesario contar con criterios específicos y fuentes objetivas de información que definan el rasgo que mide la prueba. Aiken (1995) define la validez como "el grado en el cual ésta mide aquello para lo que se diseñó".

Una prueba psicológica no tiene una sola validez, sino varias, que están en función del uso para el que ésta se destine y el grupo en el que se utilice.

La validez se estima mediante un coeficiente de correlación, denominado *coeficiente de validez*, que indica la relación existente entre los datos obtenidos de la prueba psicológica y los datos que se utilizan como índices para los puntajes del individuo en la variable de criterio. Magnusson (1973).

Brown (1980) menciona que los diversos procedimientos para conocer la validez de una prueba se clasifican en tres categorías principales según los *Standards for Educational and Psychological Tests*: **validez de contenido**, **validez relacionada con el criterio**, y **validez de construcción**.

A continuación se explicará en que consiste cada uno de estos procedimientos.

2.5.4.1. Validez de contenido.

Este tipo de validez supone un adecuado muestreo de contenido, es decir, una selección cuidadosa de los reactivos que conforman la prueba psicológica, para lo cual se realiza un examen sistemático del contenido de los reactivos que conforman el instrumento. En el muestreo de contenido se debe especificar la amplitud que fue considerada, y esta debe tratar de abarcar todas las áreas posibles. Este análisis permitirá determinar si éstos comprenden una muestra representativa del atributo o rasgo que se pretende medir.

Uno de los procedimientos para determinar si una prueba o test tiene validez de contenido consiste en que un grupo de jueces expertos en el tema decidan si los reactivos o ítems de ésta son una muestra representativa del área que se pretende medir, si los jueces deciden que la muestra es representativa, entonces el instrumento tiene validez de contenido. Para que este procedimiento pueda ser considerado objetivo es necesario especificar a los jueces el dominio de contenido que fue considerado, así como las categorías y su grado de importancia.

Un procedimiento más para conocer la validez de contenido consiste en administrar la prueba un grupo de sujetos, someter este grupo a un curso que

contenga información relativa a lo que la prueba pretende medir, y al finalizar el curso se administra nuevamente el instrumento. Este procedimiento indicará si la prueba mide la información proporcionada en el curso.

Algunos autores mencionan la *validez aparente o de facie* al describir la validez de contenido. Esta validez, es la menos objetiva, ya que sólo toma en cuenta la apariencia de la prueba, es decir, lo que aparentemente parece medir. Se considera importante que la mayoría de las pruebas tengan este tipo de validez, ya que ésta permite cierto grado de confianza en las personas que van a utilizarla. Una prueba que a simple vista no indica que es lo que pretende medir es probable que genere desconfianza y poca cooperación por parte de los sujetos. Este tipo de validez no debe confundirse con la validez de contenido, puesto que no es una validez en sentido técnico. Es importante señalar que "*la validez aparente no se debe substituir por la validez objetivamente determinada*" (Anastasi, 1988).

Para obtener la validez aparente es necesario que en el momento en que se elabore el instrumento se tenga en cuenta la población a la que va dirigida, ya que se requiere de una terminología especial para cada caso, por ejemplo, no se podrían utilizar los mismos términos para evaluar a niños y adultos o a un ingeniero y a un mecánico.

2.5.4.2. Validez empírica o de criterio.

Los tests psicológicos se emplean en algunas situaciones para predecir conductas en situaciones específicas. En este caso se requiere comparar los resultados de la prueba con un criterio específico. El *criterio* es una medida directa e independiente de lo que mide la prueba (Anastasi, 1988). Un ejemplo de esto es comparar las puntuaciones obtenidas de un individuo en un test de inteligencia para predecir sus logros académicos en la universidad. Las medidas de criterio para obtener la validez de una prueba se pueden obtener mediante dos formas, al mismo tiempo en que se aplica la prueba o después de un intervalo de tiempo.

La *validez de criterio* se puede clasificar en *validez predictiva* y *validez concurrente*, la primera se emplea como se mencionó anteriormente para predecir la actuación o ejecución de un individuo ante determinada situación, y la medida del criterio se obtiene después de un determinado intervalo de tiempo; este tipo de validez se emplea comúnmente en los instrumentos empleados en la selección y clasificación de personal. La validez concurrente se emplea cuando ya se tienen datos de los individuos (criterio) y éstos se pretenden substituir con los resultados

obtenidos en la prueba, esto es con la finalidad de ahorrar tiempo en determinados procesos, esta validez se emplea generalmente en instrumentos de diagnóstico.

Debido a que la validez empírica requiere de medidas de criterio, es necesario tener ciertas consideraciones al elegir éstas.

Brown (1980) señala tres principales características para elegir una medida de criterio: *relevancia*, *confiabilidad* y *desviaciones tendenciosas*.

La *relevancia*, considera una evaluación racional de las dimensiones relevantes del criterio conceptual para determinar si están representadas o no en la medida del criterio.

La *confiabilidad* de la medida de criterio, se refiere a que si la ejecución del criterio varía con el tiempo no se puede relacionar de manera consistente con otras medidas.

Y por último, es importante que no haya *desviaciones tendenciosas* que afecten las medidas, para evitar esto señala que si han de utilizarse medidas de criterio como juicios, será necesario proporcionar indicaciones específicas y concretas de las características que se van a calificar para que el procedimiento de calificación sea lo más objetivo posible. Brown (1980) menciona que una fuente de desviación es la contaminación de criterios que se refiere a "*la situación en que la calificación de una persona en un criterio se ve afectada por el conocimiento del evaluador de su calificación predictor*".

Concluye señalando que la mejor medida de criterio es la que tenga mayor número de ventajas prácticas, es decir, la que sea más fácil de utilizar, que se disponga de ella con facilidad y que no resulte muy costosa.

En seguida se presentan algunos de los métodos que sugieren Brown (1980) y Anastasi (1988) para obtener la validez de criterio.

El método más común para obtener la validez de criterio es el de obtener un *coeficiente de validez*, éste se obtiene mediante la correlación de las calificaciones de la prueba con un determinado criterio. Este procedimiento requiere de cinco etapas que incluyen la selección de un grupo apropiado; la aplicación de la prueba predictor; la aplicación del tratamiento relevante; la recopilación de los datos del criterio y por último se obtiene la correlación de las calificaciones y el criterio. Es importante señalar que si la correlación de las

variables es lineal se puede subestimar la magnitud de la correlación y que esta magnitud estará determinada por el rango de las diferencias individuales.

Otro método es el de *grupos contrastados o diferenciación de grupos* que implica comparar dos grupos, uno que haya sido sometido a tratamiento (ejecución en el criterio) y uno control, si se encuentran diferencias significativas entre ambos se dice que la prueba tiene validez.

La *exactitud en la toma de decisiones o desempeño en la tarea*, es un método más para conocer la validez de criterio y consiste en realizar una serie de registros acerca de la actuación en el criterio determinado y determinar la proporción de éxitos o fracasos en los sujetos que fueron seleccionados con una prueba predictora. El test psicológico que tenga mayor validez será entonces aquel que proporcione mayor número de decisiones correctas.

La *utilidad* de la prueba, test o instrumento, es un método más para determinar la validez de criterio, éste se logra mediante un análisis de costos y beneficios de diversas acciones. Los resultados del instrumento deben representar algún beneficio y ofrecer un ahorro de costos en la toma de decisiones, para ser considerada como válida.

2.5.4.3. Validez de construcción o validez estructural.

Este tipo de validez se emplea para comprobar un constructo que se supone mide la prueba o test. *"En la medida en que una variable es abstracta y latente, más que concreta y observable se le denomina constructo. Es algo que los científicos construyen y que no existe como dimensión observable de la conducta. Un constructo refleja la hipótesis de que una variedad de conductas se correlacionan entre sí, en estudios de diferencias individuales serán afectadas de manera semejante por manipulaciones experimentales"* (Nunnally, 1995).

Un ejemplo de un constructo es el estudio de la ansiedad o la personalidad, por la tanto, las pruebas o tests psicológicos que tienen validez de construcción miden rasgos para los cuales no hay un criterio externo (Magnusson, 1969) o algún atributo o cualidad que se supone poseen las personas (Brown, 1980).

Nunnally (1995) menciona que *"el objetivo del estudio de los constructos es emplear una o más medidas cuyos resultados se generalicen a una clase más amplia de medias que legítimamente empleen el mismo nombre"*.

Para obtener la validez de construcción es necesario reunir datos suficientes que permitan comprobar que los resultados del test pueden medir un constructo, Brown (1980) señala que no es posible obtener un índice cuantitativo de la validez de construcción.

Los pasos para obtener la validez de construcción son los mismos que utiliza el método científico, que incluyen formulación de hipótesis, observación del fenómeno, recopilación de datos y por último aceptar o rechazar las hipótesis.

Existen diversos métodos para obtener la validez de constructo, en seguida se presenta brevemente en que consiste cada uno de ellos, según la clasificación que presenta Brown (1980):

Método intrapruebas: Consiste en estudiar la estructura interna del test psicológico: su contenido, formas de respuesta, y las relaciones entre las subpruebas o los reactivos que lo forman. Este método considera la determinación del universo medido por el constructo y otras variables. La validez de contenido de un instrumento representa una fuente de evidencia para la validez de constructo, ya que al especificar el universo conductual se define la naturaleza del constructo que mide el test.

Método interpruebas: Este método toma en consideración las relaciones recíprocas entre varias pruebas de manera simultánea, indicando cuales son los aspectos que tienen en común éstas, o si miden o no el mismo constructo. Encontrándose en este método la validez congruente que consiste en obtener una correlación entre las calificaciones de una prueba nueva y otras ya existentes.

Y el *análisis factorial*, Nunnally (1995), define el *análisis factorial* como un "procedimiento estadístico que permite encontrar agrupamientos de variables relacionadas. Cada grupo o factor, consiste en un grupo de variables cuyos miembros se correlacionan de manera más elevada entre ellos mismos de lo que hacen con variables fuera del grupo. Cada factor se toma como un atributo que es medido en mayor o menor grado por instrumentos particulares de acuerdo con su correlación con el factor". Kline (1993), afirma que el *análisis factorial* es un método estadístico en donde las variaciones en las calificaciones en cierto número de variables están expresadas en un pequeño número de dimensiones o constructo. Algunos autores señalan además que estas correlaciones son denominadas a veces como "validez factorial" de las medidas, pero es mejor hablar de dichas correlaciones como la "composición factorial".

Método de estudios relacionados con el criterio: En este método se toma en cuenta la naturaleza y el tipo de criterio que predice un test psicológico, para dar una idea del constructo que el instrumento está midiendo. Por lo tanto, la validez de criterio puede proporcionar información valiosa para la validez de construcción.

Método de manipulación experimental: Consiste en manipular una variable y observar los efectos que esta produce en las calificaciones obtenidas en el test psicológico. La variable que se manipulará puede ser el constructo que se pretende medir. De esta manera se puede observar si los resultados obtenidos a partir de la aplicación del instrumento comprueban las hipótesis que plantea el constructo.

Método de estudios de capacidad de generalización: Consiste en analizar sistemáticamente la prueba sobre una variedad de condiciones o dimensiones o con diferentes condiciones en la administración.

Este método se basa principalmente en la contribución de Campell y Fiske(1959), (citado en Brown, 1980 y Nunnally, 1995) sobre estudios de la validez de constructo, en el cual se hace mención al método de *matriz de multirasgo y multimétodo*, el cual sostiene que hay cuatro tipos de correlaciones:

- 1) *Coefficientes de confiabilidad:* Describen el grado en que una medida es consistente internamente.
- 2) *Correlación heterorasgo-monométodo:* Denota la correlación entre dos medidas que comparten un método común pero evalúan atributos diferentes.
- 3) *Correlación monorasgo-heterométodo:* Denota una correlación entre dos medidas del mismo rasgo que utilizan métodos diferentes.
- 4) *Carrelación heterorasgo-heterométodo:* Es una correlación entre atributos diferentes que usan métodos diferentes.

Es importante señalar que los *coeficientes de confiabilidad* esperados sean elevados, mientras que las *correlaciones heterorasgo-heterométodo* deben ser más bajas puesto que utilizan métodos diferentes y miden rasgos distintos.

2.5.5. Normalización.

Un punto más a considerar en el trabajo con los tests o pruebas psicológicas, es la interpretación de los datos que éstas proporcionan, ya que la calificación aislada de una prueba sólo proporciona puntajes crudos que no se pueden interpretar en situaciones prácticas. Para lograr una adecuada interpretación es necesario hacer una normalización de las calificaciones.

La calificación que obtiene un sujeto en una prueba o test depende principalmente de su ejecución en ella, y en algunas pruebas del grado de dificultad que ésta presenta, para poder interpretar esta calificación es necesario recurrir a normas. *"Las normas son cualesquiera datos estadísticos que proporcionen un marco de referencia para interpretar los puntajes de un individuo en relación con los puntaje de otros"* (Nunnally, 1995).

Magnusson (1973) señala que *"la normalización es una transformación de la distribución de los puntajes originales a una distribución normal que permita tener escalas con diferentes niveles de dificultad"*. De ésta manera las calificaciones de una prueba podrán interpretarse con mayor precisión.

La finalidad de convertir los puntajes a una distribución normal es que ésta proporciona un significado estadístico preciso, en el cual el porcentaje de individuos que se encuentra arriba y abajo de cada puntaje se conoce exactamente cuando se tiene una media y una unidad de medidas conocidas.

Anastasi (1988) señala que los puntajes normalizados tienen una doble finalidad, indicar la situación relativa de un sujeto en relación al grupo normativo que permite valorar su ejecución con relación a otras personas; y proporcionar medidas que hagan posible la comparación directa de la actuación del sujeto en distintos instrumentos.

En relación a la determinación de normas, es importante señalar que éstas se deben obtener en una muestra representativa de la población para la que fue construida para que se puedan obtener valores estables.

Por último es importante señalar que si se han de utilizar normas ya establecidas de una determinada prueba, es necesario considerar las características de la muestra para las que éstas se obtuvieron.

2.6. Clasificación de los tests.

Morales (1975) señala que los tests psicológicos pueden ser clasificados tomando en cuenta sus características, las cuales pueden ser:

- a) su objetivo,
- b) tipos de respuesta,
- c) formas de aplicación,
- d) ejecución,
- e) instrucciones,
- f) material y
- g) calificación.

Brown (1980) hace una clasificación de acuerdo al formato de cada prueba. Se expone en seguida una clasificación considerando a ambos autores.

a) En principio los tests psicológicos se clasifican por su *objetivo*, es decir, el objetivo principal para el que fueron construidos y los rasgos o conductas que pretenden medir. Esta clasificación divide a las pruebas en:

- 1) *pruebas de potencia o ejecución máxima y,*
- 2) *pruebas de rasgos o ejecución típica, que también incluirían a las técnicas proyectivas.*

Una característica de las pruebas de *ejecución máxima*, es que éstas presentan estímulos y tareas específicas al sujeto, mientras que en las de *ejecución típica*, que incluyen a las técnicas proyectivas los estímulos y tareas que se le piden al sujeto son ambiguos.

1) Las *pruebas de ejecución máxima o de potencia* son aquellas en las cuales se evalúa el rendimiento y las aptitudes del individuo, estas pruebas demandan el mayor rendimiento del sujeto en la tarea encomendada. La calificación que obtiene el individuo está determinada por tres factores principales: *habilidades*

innatas, habilidades aprendidas y la motivación que presente al resolver la prueba.

Dentro de esta clasificación se encuentran las pruebas de inteligencia, rendimiento, aptitudes múltiples y específicas. Las pruebas de rendimiento proporcionan información acerca de los conocimientos, habilidades y capacidades del individuo. Las pruebas de aptitudes permiten hacer una predicción de la conducta futura que tendrá el sujeto en determinadas tareas.

2) Las *pruebas de ejecución típica o de rasgos* permiten conocer la forma en que se comporta un individuo de manera habitual, la forma en que reacciona, piensa y siente ante diversas situaciones; estas pruebas se pueden clasificar además en estructuradas e inestructuradas.

En esta clasificación se encuentran los inventarios de personalidad, actitudes, intereses, preferencias, valores y las técnicas proyectivas.

b) Las pruebas se pueden clasificar por su *forma de respuesta* en:

1) *objetivas* y

2) *subjetivas*.

En las *objetivas* se encuentran las pruebas de inteligencia, y en las *subjetivas* las técnicas proyectivas.

En esta clasificación se puede considerar además las formas que se presentan al sujeto para elegir su respuesta. Una prueba, puede entonces, presentar respuestas alternativas o respuestas libres, en las primeras el sujeto deberá elegir su respuesta de entre varias opciones, y en las de respuesta libre el sujeto deberá completar frases, hacer historias, realizar un dibujo, etc.

c) La clasificación de acuerdo a la *forma en que se administra* una prueba señala tres formas para aplicar una prueba:

1) *individual*,

2) *colectiva* y,

3) *autoadministrable*.

Las pruebas de *aplicación individual*, requieren la presencia de un examinador para cada sujeto, en este tipo de aplicación el examinador indicará al sujeto las tareas que habrá de realizar y manipulará los materiales que sean necesarios. Las pruebas de *aplicación colectiva* son aquellas que se pueden aplicar a más de un sujeto a la vez, generalmente son del tipo lápiz-papel, y no requieren manipulación de materiales, la tarea del examinador será dar las instrucciones, vigilar a los sujetos y atender las dudas que pudieran presentar. En las *pruebas autoadministrables*, la tarea del examinador será sólo dar las instrucciones de cómo resolver la prueba, pero no será necesario que éste permanezca mientras el sujeto resuelve la prueba.

Dentro de las formas de aplicación cabe señalar, una de las más recientes, a la cual se le denomina *automatización*, esta forma de aplicación requiere de una computadora en la cual se le presentarán los reactivos al sujeto, para posteriormente ser calificadas automáticamente. (Murphy, 1994).

d) Las pruebas psicológicas se clasifican de acuerdo a *la manera en que se pueden contestar* en:

1) *pruebas de poder y*

2) *pruebas de velocidad.*

Las *pruebas de poder* contienen reactivos de diferentes grados de dificultad, y tienen un tiempo límite suficiente para poder contestar todos los reactivos. Las *pruebas de velocidad* generalmente contienen reactivos de menor dificultad, y tienen un tiempo límite en el cual el sujeto deberá contestar el mayor número de reactivos que le sea posible.

e) Las *instrucciones* que puede contener una prueba psicológica pueden estar escritas en el protocolo, o estar especificadas por el examinador de manera oral, en algunas ocasiones será necesario que el examinador ejemplifique cómo ha de realizarse la tarea.

f) Las pruebas psicológicas se pueden clasificar de acuerdo *al material que empleen* en:

1) *lápiz-papel,*

2) *verbal,*

3) *ejecución,*

4) *aparatos específicos para realizar alguna tarea y,*

5) *una combinación de los tres primeros.*

g) La *calificación*, de las pruebas se pueden llevar a cabo manualmente o por medio de una computadora, representando este último método un ahorro de tiempo y una mayor objetividad en las puntuaciones de cada sujeto.

h) Por último, es importante señalar que las pruebas psicológicas deben tener un *límite de aplicación*, éste considera la población para la que se elaboró la prueba. De esta manera se clasifican en pruebas para bebés, niños preescolares, niños escolares, adolescentes y adultos.

2.7 Influencia del examinador en cuanto a la aplicación.

Un test psicológico proporciona información útil para la toma de decisiones en diversas actividades, tales como, elaborar un diagnóstico, determinar un tratamiento, elegir la carrera que ha de estudiarse, o aceptar o no a una persona para determinado empleo, y como es de esperarse, estas decisiones pueden influir o determinar de alguna manera el futuro de las personas, por lo tanto, los resultados que brinda un test son información muy delicada, que requiere de un manejo profesional, es decir por personal debidamente capacitado para aplicar, calificar e interpretar los resultados.

En seguida se presentarán algunas de las consideraciones más importantes que señalan diversos autores (Morgan, 1975; Anastasi, 1988; Sattler, 1988; Aiken, 1995) que se deben tener en cuenta tanto para el personal que ha de administrar la prueba (examinador), así como las condiciones que se deben tomar en cuenta durante el proceso de aplicación.

El examinador o aplicador de una prueba debe cubrir determinadas características y habilidades para poder desempeñar profesionalmente su labor. En principio debe contar con conocimientos suficientes acerca del comportamiento humano y de las diferencias individuales que presenta cada sujeto, esto con el propósito de conocer la importancia del trabajo que ha de realizar, entender a los diferentes sujetos y poder establecer una relación interpersonal adecuada. Las personas que pueden desempeñar esta labor pueden ser psicómetras o psicólogos con preparación psicométrica, es decir, personas que

cuenten con los conocimientos relacionados con el manejo de pruebas, material, instrucciones y forma en que ha de aplicarse cada instrumento.

Morgan (1975) hace algunas sugerencias de las habilidades que requiere el entrevistador completo, que pueden aplicarse al psicólogo que hace uso de los tests, *estas habilidades son sociales, de comunicación, analíticas e interpretativas y para la toma de decisiones.*

Las *habilidades sociales* le permiten al examinador mantener una relación interpersonal cordial y amable, para de esta manera obtener mayor cooperación y participación de los sujetos, además le permitirán determinar que es debido o no hacer en las diferentes situaciones que puede afrontar.

Las *habilidades de comunicación* le han de servir para mantener una correcta interacción verbal, que incluye el poder determinar el vocabulario que ha de emplear según la edad, escolaridad y educación de los sujetos. El aplicador deberá saber modular los tonos de voz que ha de utilizar, y ser claro en sus explicaciones para que no existan problemas en la forma de resolver la prueba.

Las *habilidades analíticas e interpretativas* le serán de utilidad para poder ser objetivo en el momento de calificar e interpretar las puntuaciones de cada sujeto obtenidas en las pruebas psicológicas.

Las *habilidades en la toma de decisiones* le servirán para no dejarse llevar por estereotipos al elaborar los reportes finales con los que se han de tomar decisiones, es decir, que no deberá juzgar a las personas por su sexo, edad, raza o religión.

Las condiciones para poder administrar adecuadamente un test psicológico requieren de un lugar amplio y bien iluminado, en el que no haya exceso o falta de luz, con ventilación adecuada y temperatura agradable, deberá contar con una silla cómoda y una mesa para cada sujeto, así como el que este aislado de ruidos y de todo tipo de interrupciones que puedan distraer la atención del sujeto.

En el proceso de administración, el examinador necesita establecer un adecuado *rappor*t, es decir, deberá realizar un esfuerzo para despertar el interés del sujeto, obtener su cooperación y asegurar que sigue las instrucciones tipificadas de la prueba (Anastasi, 1988).

Para poder establecer el *rapport*, es necesario que haya empatía con el sujeto, y para lograrla, el examinador deberá empezar el proceso de administración saludando al (los) sujeto (s), presentándose, preguntado su nombre al sujeto, en el caso de que la aplicación fuera individual. En esta parte el examinador se puede ayudar con expresiones faciales y modulaciones de voz apropiadas que le permitirán establecer un ambiente de confianza.

Anastasi (1988), señala que cualquier test representa una amenaza implícita al prestigio del individuo, por lo que será necesario ganarse la confianza y cooperación del sujeto, así como darle cierta seguridad al iniciar la prueba. Sattler (1988) menciona que para reducir la angustia que pudiera tener el sujeto, se le debe explicar para que se le está aplicando la prueba y como afectará su futuro.

Por último se presentarán los puntos principales de la lista de sugerencias que propone Sattler (1988) para el examinador, en el proceso de aplicación de pruebas de inteligencia, tomando únicamente aquellos puntos que se pueden extender a la mayoría de las pruebas psicológicas.

I.- Precauciones generales.

A.- Lea, aprenda y vuelva a leer las instrucciones.

B.- Atégase siempre a los procedimientos estandarizados.

1. Use la redacción exacta.
2. Observe con mucho cuidado el tiempo establecido.
3. Presente los materiales en la forma descrita.
4. Siga con todo rigor las instrucciones para calificar.
5. No se conforme con leer las instrucciones impresas, téngalas a la mano para consultarlas.

C.- Sea objetivo.

1. No haga ninguna insinuación sobre lo correcto o incorrecto de las respuestas del sujeto.
2. No de ninguna pista sobre la respuesta que espera; cuide mucho su tono de voz, recuerde que está en examen, no enseñando.

D.- Sea natural.

1. Trate de ser impersonal, pero sea amable.
2. Aprenda a usar una dicción estandarizada en forma natural e informal.

3. Establezca empatía recíproca y comunicación franca antes de que inicie la administración; tome una actitud expectante.

E.- Prepare el ambiente

1. Evite distracciones.
 - a) Visuales: Impida que el sujeto mire hacia puertas y ventanas donde haya mucho movimiento y actividad. Que no vea tampoco hacia grandes espacios abiertos que tengan imágenes, colores, etc. que puedan distraerlo. Evite el ruido.
 - b) Auditivas: Evitense las áreas ruidosas, otras voces, etc.; aplique la prueba en un lugar aislado.
 - c) Emocionales: No administre la prueba cuando el sujeto se sienta presionado, perturbado o enfermo.

II.- Administración y calificación.

F.- Eficiencia para la forma en que se administre la prueba y el manejo de materiales.

1. Establezca un orden eficiente y un método apto para manipular los materiales.
 - a) Para registrar.
 - b) Para consultar el manual *sin que se convierta en obstáculo entre usted y el examinado.*
 - c) Para guardar y sacar en forma oportuna los materiales.
2. Haga una transición suave de una prueba a otra y de un reactivo a otro. En cada punto de la prueba debe saberse qué es lo que se va a presentar en seguida.
3. Conozca el material, las calificaciones lo suficiente como para no alargar la prueba en forma innecesaria. Una prueba demasiado prolongada suele producir cansancio, desinterés o ambos.
 - a) Disimule o reoriente los comentarios no importantes.
 - b) Reduzca al mínimo los movimientos extraños fomentando el interés, la motivación, la orientación de las tareas.
 - c) Prevenga el cansancio y la distracción.

G.- Calificación.

1. Es esencial que el examinador conozca bien los criterios de calificación. Este requisito se refiere en especial a la comprensión del *objetivo* de cada prueba o subprueba.
2. Debe recordarse que los criterios de calificación son simplemente lo que dice la "etiqueta". "Criterios" y no respuestas completas y de absoluta veracidad o falsedad. Suele ser necesario evaluar respuestas equivalentes basándose en las enumeradas en los criterios de calificación, ya que no pueden incluirse todas las respuestas posibles.
3. El examinador principiante debe constatar las respuestas con el manual, para verificar cualquier duda.
4. Deben volver a revisarse cada uno de los pasos en el proceso de calificación.

H.- Cuidado del material de la prueba.

1. Siempre que cualquier material presentado al niño tenga marcas o esté deteriorado de modo que pudiera influir en su respuesta, habrá que reemplazarlo.
2. Los materiales extravagantes o dañados tendrán que ser substituidos por otros *idénticos a los originales*.

2.8. PUNTOS A CONSIDERAR SOBRE LA ADAPTACION Y ESTANDARIZACION DE INSTRUMENTOS.

Quando se planea llevar a cabo la adaptación de un instrumento de origen extranjero, o bien, una investigación de tipo transcultural, es necesario considerar numerosas dificultades antes de iniciar formalmente el estudio. Varios investigadores han intentado aplicar mediciones de la personalidad en diferentes culturas, o bien, han tratado de adaptar instrumentos de una cultura a otra; sin embargo, su contribución a las teorías de la personalidad modal o estructura típica de personalidad nacional esta oscurecida, debido a los problemas metodológicos circundantes y a las limitaciones en los instrumentos.

Muchos teóricos han encontrado para su propia satisfacción características universales, sin contemplar, el origen cultural. Subsecuentemente estos expertos han intentado encontrar un desarrollo universal u otros procesos psicológicos, que no han podido satisfacerse debido a serios problemas metodológicos relacionados con la equivalencia de las mediciones. Actualmente, la universalidad de la conducta en general ha sido ubicada en la idea de que existe un proceso básico en la estructura de la personalidad sin contemplar las experiencias culturales, la equivalencia que se encuentra opuesta entre los diferentes expertos de la psicología que trabajan a nivel transcultural, o bien, en la adaptación de instrumentos para su uso en psicología aplicada usualmente no es tan aparente; no obstante, conceptos como *dominancia o sumisión* que son definidos en un escenario cultural particular con ciertas características explícitas, no tienen un significado psicológico equivalente en una segunda cultura, a pesar de que la cultura externada es similar, es dudoso que la estructura y el contenido de la personalidad sean las mismas en todos los países, y en las diferentes culturas (Butcher y Pancheri, 1976).

Así mismo, es necesario resaltar que hasta en grupos nacionales altamente similares los conceptos psicológicos deben ser examinados con cuidado para determinar los vínculos de equivalencia y diferencia.

Entender una conducta dentro del contexto de una cultura diferente es frecuentemente problemático, debido a que esta puede mostrar un diferente origen o significado en las diversas escenografías culturales. Por lo tanto, los expertos han recomendado llevar a cabo observaciones sistemáticas de la conducta dentro del ámbito cultural en la cual ocurre.

Evidentemente, bajo el sustento transcultural es menester efectuar contrastaciones generales, a menudo entre diferentes agrupaciones culturales. Para proceder con el aspecto empírico que marca el método transcultural, se tendrán que comparar diferentes culturas (países) con mediciones semejantes o medidas equivalentes generales con apoyo diversas estrategias.

Las mediciones provisionales y pesadas frecuentemente aplicadas en los estudios con aspiraciones transculturales producen resultados curvados, ya que los investigadores utilizan los productos que se ofrecen en el mercado en vez de diseñar instrumentos de medición apropiados para una determinada situación.

Este tipo de instrumentos o pruebas presentan graves problemas, debido a que el significado de los resultados son altamente cuestionables. Este hecho

arrojará respuestas poco comunes que no pueden ser compartidas con los datos provenientes de otros grupos de diversas culturas. Los estudios realizados se inician usualmente con la extracción de una muestra de sujetos, posteriormente se seleccionarán los instrumentos y con el apoyo de la herramienta estadística se intentará efectuar la generalización al total de la población. El método exige que se utilice la técnica conocida como *selección aleatoria*; esto con el propósito de que todos los sujetos de la población tengan la misma oportunidad de integrar la muestra.

Para conducir una investigación completa y generalizable en el ámbito transcultural y nacional o cualquier estudio que intente la adaptación de instrumentos de origen extranjero con estas aspiraciones, requiere el contar con un conocimiento íntimo de la población en estudio, instrumentos de medición, método de investigación, así como la situación política, social y económica actual. Así mismo, es importante contar con la ayuda de psicólogos nativos del país receptor a fin de planear las estrategias de investigación, guiar la recolección de los datos, solicitar la participación de los expertos nacionales durante todas las fases del estudio y utilizar instrumentos estandarizados, pero estos no pueden ser empleados para todos los países o culturas, cuando están redactados en un formato incompatible. Para llevar a cabo la tarea de adaptar un inventario de personalidad, o bien, de cualquier instrumento de origen extranjero a una nueva lengua y cultura, Butcher y Pancheri (1976), recomendaron cumplir con los siguientes cuatro lineamientos:

1.- Indagar acerca de la existencia de variables a medir en la población receptora, pero no se trata de determinar únicamente la presencia, sino demostrar que ciertas características de personalidad presentan una distribución semejante dentro de la población.

2.- La labor tiene que ser relevante y significativa para la nueva población; esto se refiere a que el instrumento tanto verbal como no verbal tiene que ser apropiado para la nueva población y contener el mismo significado en ambas o más poblaciones/países.

Es necesario conducir estudios preliminares y llevar a cabo modificaciones para obtener un formato apropiado y esto también se puede prevenir, efectuando la traducción del instrumento mediante una estrecha colaboración de los constructores de formatos del país receptor.

3.- El atributo más importante de los inventarios de personalidad, y de los tests en general es la validez. Un instrumento sin validez demostrada a nivel empírico, no permite tomar ninguna decisión respecto a los resultados obtenidos. Existen graves problemas cuando una prueba o test traducida textualmente se ofrece en el mercado sin haberla validado y el usuario piensa que los errores de la medición establecidos en el país de origen pueden extenderse al país receptor. La labor de demostrar la validez de los instrumentos traducidos es responsabilidad del traductor, que debe ser compartida con la casa editorial.

Es importante señalar que no existe un monitoreo o parámetros éticos para la traducción de instrumentos en México. Esto trae como consecuencia errores graves en la toma de decisiones a partir de los datos obtenidos con instrumentos no adaptados, mal adaptados, etc.

4.- Tanto los traductores como los editores asumen la misma responsabilidad en el seguimiento de la traducción. Esto sugiere que el problema del control de la educación de la traducción y diseminación del instrumento no es simple y esto únicamente ha sido reportado por la American Psychological Association APA (Asociación Americana de Psicología).

La traducción de los reactivos (ítems) es probablemente la tarea más complicada y el aspecto más importante de la investigación con instrumentos de origen extranjero, en el terreno de la instrumentación en la psicometría. Es esencial obtener una traducción adecuada para asegurar la equivalencia de la medición. Así mismo, los problemas de traducción y adaptación nunca pueden ser solventados satisfactoriamente para prever mediciones equivalentes, se sugiere que los expertos del área deban construir reactivos comparables, para ser utilizados en las diversas culturas o países; siendo necesario resaltar, que en la traducción de los instrumentos de medición no es posible obtener una equivalencia genuina. Así mismo, a pesar de obtener una versión lingüística perfecta de los instrumentos, esto no asegura respuestas similares en los probandos, debido a que las palabras que integran los reactivos no poseen la misma frecuencia en ambas lenguas. Entre mayor sea la distancia entre una pareja de culturas, mayores serán los problemas de la traducción y adaptación.

Es importante distinguir entre la traducción de instrumentos objetivos de características de personalidad para el uso habitual y los que se construyen para realizar investigaciones.

Para el uso habitual, la traducción debe ir dirigida a eliminar algunos reactivos de la versión original y proveer sustitutos y este es un eslabón necesario en la validación de la nueva versión. Este procedimiento, sin embargo, limita la medición para comparaciones transculturales ya que los reactivos eliminados son precisamente aquellos que pueden diferenciar a las dos culturas. Este fenómeno se conoce como *equivalencia paradójica* y para salvar este problema, los dos autores referidos, recomendaron que los instrumentos traducidos deben someterse a los usuales procedimientos de validación, mediante las cuales, algunos reactivos pueden eliminarse según el peso que muestren.

Desarrollar una traducción precisa es esencial para maximizar la equivalencia de las mediciones de las características de la personalidad.

2.8.1 Problemas con los que debe enfrentarse el traductor

1.- Los traductores frecuentemente se enfrentan con el problema de elegir entre el empleo de una expresión rara o sofisticada, que es una equivalente exacta a la original, pero que presta una expresión equivalente deficiente y el utilizar una forma expresiva más familiar, que hace que el texto sea comprensible y aceptable, por ello se recomienda utilizar simples oraciones verbales.

2.- Se recomienda emplear términos concretos y no abstractos, ya que las traducciones muy generales son sumamente difíciles de comprender y por lo tanto, se debe favorecer y exigir la especificidad en la traducción de instrumentos para la medición de características de personalidad.

3.- Algunas expresiones verbales pueden tener limitaciones, ya que en una determinada cultura pueden ser expresadas de manera figurativa, en otra en cambio literal; en estos casos, se sugiere agregar una explicación anexa al reactivo, especificando la intención de la versión original, esto permitirá que el reactivo muestre una mayor equivalencia.

4.- Las expresiones idiomáticas en la lengua original no siempre tienen equivalentes en la lengua receptora, para cubrir la precisión del significado, el traductor de instrumentos podría usar ambas; la traducción idiomática y literal del reactivo, manteniendo la última forma entre paréntesis.

5.- El tiempo del verbo frecuentemente ocasiona dificultades en la traducción. En muchas lenguas el tiempo y el género son utilizadas con más precisión que en inglés; por lo tanto, el empleo de cierto tiempo en la traducción, no corresponde al mismo significado que en el texto original, para resolver esta

situación se recomienda utilizar dos tiempos en la versión traducida para cubrir el significado exacto del texto original.

6.- En muchas lenguas las expresiones dobles negativas son usadas cuando en inglés únicamente se presenta la negativa singular; esto origina un grave problema en los inventarios de personalidad con un formato de respuestas dicotómicas (verdadero-falso), ante esta situación, se recomienda que los reactivos problemáticos sean aclarados adicionando una explicación a cada reactivo que produce ese efecto de confusión en la doble negatividad. Una solución alterna es modificar el reactivo de la forma negativa a la positiva, con el cambio respectivo de la dirección del puntaje (calificación) del reactivo.

7.- Para asegurar el éxito de la traducción y cubrir con precisión el significado original en la lengua receptora, se requiere que los traductores muestren un elevado grado de familiaridad en ambas lenguas.

La utilidad de la traducción de los instrumentos dependerá de la extensión de la inclusión de términos que se refieren a las experiencias reales, es decir, que sean familiares en ambas culturas. Cuando dos culturas se encuentran muy distanciadas, será imposible cubrir el aspecto de la equivalencia del significado para algunas palabras o expresiones. Se sugiere que estos reactivos problemáticos sean examinados, sean eliminados o reemplazados por un mejor significado equivalente, que puede ser determinado problemáticamente por una estructura factorial similar.

Se han desarrollado múltiples procedimientos para la traducción de reactivos con la meta de ofrecer una elevada equivalencia en la traducción.

El procedimiento más común y el menos adecuado es el de *traducción directa*, en este tipo de aproximación, un traductor bilingüe traduce el material de una lengua a la otra lo mejor que puede. Este procedimiento se encuentra limitado por la experiencia, antecedentes culturales y habilidad lingüística del traductor. En este método, especialmente cuando es realizado por una sola persona dedicada a esto, la traducción se torna inadecuada.

Algunos expertos sugirieron formar un equipo o comité de abordaje para realizar una traducción directa. Esto es, un equipo de bilingües traducirán de una lengua a la otra en forma independiente, pero colaborando para producir la mejor traducción compuesta, que permite controlar el sesgo individual. No obstante, este proceso muestra una deficiencia, debido a que los integrantes del comité

frecuentemente titubean o se inclinan a criticarse en forma negligente entre ellos. Así mismo, se ha observado que existen pocos expertos suficientemente familiarizados con las dos lenguas y las dos culturas asociadas; por lo tanto, el mejor acercamiento para lograr una traducción adecuada es el *método de la colaboración*, proponiéndose el utilizar dichas traducciones a psicólogos bilingües expertos en psicometría de ambos países para que trabajen sobre las mejores equivalencias lingüísticas y psicológicas.

El método de traducir el contenido verbal que reduce el error y el sesgo es la *traducción hacia atrás o retrospectiva*. Mediante este procedimiento, uno o más bilingües traducen el material verbal de la lengua original a la receptora. El material de traducción se vuelve a traducir hacia atrás, es decir, hacia la lengua fuente por bilingües diferentes.

Las dos versiones (original y forma traducida) son comparadas y se examinan las discrepancias, si existen reactivos que en las dos versiones presentaron diferencias en el significado, se repite la traducción. Esta técnica puede considerarse como un procedimiento idóneo para verificar la precisión lingüística; no obstante, este método tampoco permite empíricamente los no equivalentes.

Los traductores deben compartir una serie de reglas para traducir oraciones buenas a partir de escritos pobres de la versión receptora en la traducción hacia atrás. La estructura gramatical de la lengua origen frecuentemente es conservada durante la traducción, facilitando la traducción retrospectiva de la versión receptora y esto se encuentra asociado a una pobreza en la estructuración de las oraciones de la lengua receptora.

De lo anterior, se deriva que la traducción hacia atrás tampoco elimina todos los problemas de la traducción ya que contiene sistemas internos de control capaces de restringir la aparición de discrepancias, que también pueden presentarse cuando la estructura lingüística original es pobre. Otra fuente que afecta a la equivalencia puede ubicarse en aquellas palabras, frases o conceptos adecuados en la versión original que no existen en la lengua receptora.

Para linar las asperezas en la traducción, se ha sugerido el empleo de sucesivas traducciones hacia atrás que podrían favorecer mejores aproximaciones hacia el texto original. Un paso esencial en la traducción de un inventario de personalidad, sería la exploración completa de las propiedades de los reactivos y establecer cómo operan las escalas. Se ha sugerido que las frecuencias

confirmatorias sean empleadas como un examen empírico de la traducción; esto es, las respuestas de verdadero de cada reactivo en la población/país receptor se podrían contrastar con las obtenidas en una muestra relevante del grupo fuente, y los reactivos que mostraron confirmaciones externas se deberían someter a estudios especiales para establecer si las diferencias encontradas son consecuencia de las deficiencias en la traducción del instrumento.

Otros factores, fuera del contexto de la traducción, que pueden ocasionar discrepancia en las frecuencias confirmatorias pueden ser las diferencias interculturales actuales y las modificaciones del contenido de los reactivos que no permitirán la comparación. También se debe exponer, que a pesar de que los porcentajes confirmatorios de un determinado reactivo sean similares en ambas culturas/países, esto no indica que exista la seguridad que los reactivos midan efectivamente lo mismo en ambas culturas/países. Si el traductor desea alejarse de las evaluaciones subjetivas, los procedimientos analítico-factoriales podrían proveer una información evaluable tanto a nivel de los reactivos como de las escalas.

El Análisis Factorial es la herramienta que permite establecer los rangos de eficiencia de la equivalencia al grado de similitud de las estructuras factoriales del instrumento entre ambas culturas/países y los reactivos que no carguen en la estructura factorial donde deberían ubicarse en la población receptiva, podrán someterse a un examen respecto a las diferencias del significado. Este método tiene especial relevancia para los instrumentos o escalas que derivaron los métodos racionales o analíticos factoriales de la población/país fuente, ahora bien, a pesar de las bondades que ofrece el análisis factorial dentro del contexto de la instrumentación transcultural o en la investigación en psicometría, es necesario tener cuidado por el efecto de una sobre interpretación.

Otro procedimiento para eliminar lo inadecuado de la traducción de los instrumentos de personalidad es la *técnica de retest bilingüe*, ambas versiones de los instrumentos -(original y receptor)- son sometidos a un grupo de traductores bilingües, que están familiarizados con ambas culturas.

Los datos de un instrumento examinado por el test-retest bilingüe pueden ser empleados para contrastar diferencias en los reactivos y las escalas o para comparar el grado de similitud en los perfiles. Así mismo, las discrepancias encontradas a nivel de los reactivos, facilitarán orientar las modificaciones o eliminar los reactivos del instrumento.

Butcher y Pancheri (1976) ofrecieron una serie de sugerencias para realizar traducciones bilingües:

- Es necesario contar con una traducción inicial cuidadosa.
- La traducción deberá ser examinada por dos grupos de asesores bilingües seleccionados en forma aleatoria.
- Los reactivos que denoten discrepancias en las frecuencias confirmatorias deberán someterse a un escrutinio respecto a las asperezas de la traducción; las traducciones deberán enfocarse a homogeneizar la comparación de las frecuencias.
- Los reactivos que persisten en mostrar discordancias en la comparación de los porcentajes confirmatorios deberán ser eliminados o totalmente sustituidos.
- Las ventajas de este tipo de análisis únicamente se pueden alcanzarse cuando los traductores presentan capacidad fluida en ambas lenguas, así como conocimientos respecto a las dos culturas, debido a que se ha observado que proceden simplemente los antecedentes de ambas lenguas, no indica que el traductor piense en forma semejante en las dos lenguas.
- Se ha teorizado que el contexto en el cual las lenguas fueron adquiridas afectará su independencia entre ellos y por esta situación los expertos han externado que la experiencia bicultural tiende a realizar la verdadera bilingüidad. Para examinar empíricamente lo adecuado de la traducción de un instrumento de personalidad es decir, la equivalencia, se puede emplear la *correlación entre muestras bilingües*; no obstante, esta herramienta estadística no presenta la fuerza suficiente para ser concluyente respecto a la calidad de la traducción.

Una vez que el inventario de personalidad fue traducido cuidadosamente, requiere llevar a cabo la tarea de reestandarización. En esta fase de adaptación del cuestionario, si se desea realizar en forma adecuada, contempla una serie de estudios a determinar si el instrumento opera en la nueva lengua y cultura en forma semejante que en la cultura original. Se debe de subrayar que la traducción cuidadosa no es suficiente para asegurar que el trasplante del instrumento fue exitoso, para verificar la adaptación del instrumento traducido es necesario llevar a cabo una serie de estudios que fueron aglutinados bajo cinco dimensiones:

Patrones confirmatorios de los reactivos: establecer si existe una sobreposición entre los porcentajes confirmatorios de los reactivos de la versión

traducida y la original, así como demostrar la precisión de la traducción de los reactivos.

Asimismo, determinar si aparecen reactivos culturalmente sensibles o reactivos con diferencias en las frecuencias confirmatorias, no atribuibles a la traducción. En caso afirmativo, es necesario medir cómo afecto entre los resultados arrojados por las escalas o las interpretaciones.

El segundo factor se refiere a la *equivalencia de la estructura factorial* y se debe de demostrar si las escalas o reactivos mantienen una estructura de asociación semejante a través de las culturas/paises.

El siguiente factor está orientado a demostrar la *generalización transnacional de la construcción de instrumentos de personalidad*: el traductor debe iniciar su trabajo con la hipótesis que el sustento teórico del inventario de personalidad (características psicológicas), se encuentra en ambas culturas y los puntajes que miden presentan una distribución semejante; esto tiene que demostrarse en forma empírica en la fase de adaptación.

Este factor está dedicado a verificar las normas establecidas y las indicaciones para renormar, es necesario determinar si los diversos elementos tipificadores de ciertas características de personalidad, presentan semejanza entre las medidas de tendencia central y de dispersión, en caso contrario se requiere cuantificar la diferencia, así como indagar sobre las posibles estrategias homogenizadoras para hacer comparables las características de personalidad a nivel transcultural.

El último factor contempla la *validez del inventario de personalidad*: se requiere empíricamente, a través de las técnicas para obtener validez si el instrumento traducido mide lo que pretende medir en la población/pais receptor.

A pesar de los problemas en que se ha incurrido en la aplicación de los tests y a pesar de los numerosos ataques en su contra, el método de los tests se ha desarrollado ampliamente, multiplicándose y perfeccionándose las técnicas y creándose toda una rama de la estadística para favorecer su interpretación.

Todo lo anterior permite tomar en cuenta todas las características con las que debe contar un instrumento, así como los puntos a considerar cuando se utilizan estas herramientas en el quehacer diario del psicólogo, incluyendo las reglas a considerar para el análisis de instrumentos de origen extranjero.

Capítulo III

EL CORNELL INDEX

3.1 Ficha Técnica

Nombre original: *Cornell Index*

Autor: *Arthur Weider, Ph. D.*

Procedencia: *En un inicio editada por Psychological Corporation, actualmente distribuida de manera privada por el autor (USA).*

Utilidad: *Diagnostica de manera rápida síntomas psicósomáticos y neuropsiquiátricos.*

Aplicación: *Sujetos mayores de 16 años.*

Material: *Protocolo de 101 Preguntas que incluye espacio para respuestas, y plantillas para calificación.*

Formato: *Lápiz y Papel.*

Tiempo de Administración: *15 minutos.*

Normas: *Las que presenta en el manual en su versión original, tablas de rangos percentilares (ver Tabla D, de este capítulo).*

3.2. Origen del Instrumento

Durante la Primera y Segunda Guerra Mundiales las fuerzas armadas de los Estados Unidos tuvieron la necesidad de crear instrumentos de rápida administración y calificación para la selección del personal militar, ejemplos de esto son las pruebas de inteligencia *Army Alpha* y *Army Beta*, la primera verbal y la segunda no verbal, y que hoy en día de manera modificada se siguen utilizando; en lo referente a la personalidad se crearon la *Hoja de Datos Personales de Woodworth* (Personal Data Sheet of Woodworth), que constaba de 116 reactivos que evaluaban síntomas neuróticos y psicóticos, posteriormente la *Escala de Personalidad de Thurstone* (Personality Schedule of Thurstone) que estaba integrada por más de 200 reactivos tomadas de la Hoja de Woodworth, el Laird y otros cuestionarios, y que fue el la más usada durante la primera guerra mundial. Ahora bien, los avances en la detección de desordenes emocionales se

realizaron durante los últimos años de la Segunda Guerra Mundial. Enfrentados con el problema de detectar de manera masiva neuróticos potenciales, los servicios psicológicos del ejercito norteamericano diseñaron nuevos instrumentos de tipo exploratorio algunos de ellos son *Neuropsychiatric Screening Adjunct*, el *Inventario Shipley*, el *Maudsley Medical Questionnaire*, y los *Cornell Index*. Estos cuestionarios en su mayoría daban una puntuación global de ajuste personal o estabilidad emocional.

Los *Cornell Index* constaban de tres formas: dos de ellas son el *Cornell Select Index*, forma N, para uso civil, aplicable a hombres y mujeres de entre 18 y 40 años, y el *Cornell Service Index* que incluye alrededor de 30 reactivos en relación directa con el comportamiento del individuo en el servicio militar.

Al finalizar la guerra los reactivos específicamente relacionados con aspectos militares fueron modificados a fin de ser aplicables a la población civil. Estos reactivos fueron validados sobre grupos civiles, los que tenían mayor significación clínica fueron combinadas con las de la forma N, para lograr un cuestionario Cornell Index más eficaz. Esta última versión es lo que se conoce como Cornell Index FN2, contiene 64 reactivos de la forma original N más 36 adicionales y uno introductorio, dando un total de 101 reactivos (Mittelman, 1946; Cerda, 1981).

3.3. Propósito del Instrumento

El Cornell Index es un cuestionario que surgió a raíz de la necesidad de disponer de un instrumento que tuviese la capacidad de hacer una evaluación psiquiátrica y psicósomática rápida en un gran número de personas, sobre todo en situaciones en donde un examen individual resulta impracticable, al menos para la mayoría de los sujetos. El Cornell Index FN2 consta de 101 reactivos: 100 que conforman el instrumento en si, y un reactivo introductorio (el número 1), que no se toma en cuenta dentro de la calificación. El instrumento toma de 10 a 15 minutos para su contestación, y la calificación toma menos de 3 minutos utilizando las plantillas diseñadas para esta tarea.

El instrumento explora los siguientes aspectos (Weider, 1975):

1) Deficiencias de adaptación que son expresadas por sentimientos de miedo e insuficiencia.-

Se refiere a la deficiencia en la adaptación que experimenta el sujeto como sentimientos de miedo e insuficiencia, es la ineficacia que presenta el individuo como una respuesta afectiva a un peligro externo, real y actual, en donde el peligro más frecuente es la amenaza a la integridad física de la persona. Contempla los reactivos 2 al 19.

2) Reacciones patológicas del humor, especialmente la depresión.-

Es un tono afectivo de tristeza que puede ir desde el abatimiento ligero o sentimiento de indiferencia hasta una desesperación en grado extremo. Abarca los reactivos 20-26.

3) Nerviosismo y Ansiedad.-

Es el estado de tensión que revela la posibilidad de un desastre inminente, como una señal de peligro que surge de la presión de actitudes internas inaceptables que aparecen de pronto en la consciencia o bien en forma de acción, con las consecuentes respuestas de la estructura de personalidad del individuo o de la sociedad a dicha reacción. Reactivos 27 al 33.

4) Síntomas psicósomáticos neuroculatorios.-

Síndrome que se caracteriza por disnea, factibilidad aumentada, palpaciones o percepción de los latidos cardiacos, temblores, desmayos, vértigos, cefalalgia y miedo al esfuerzo. Reactivos del 34 al 38.

5) Reacciones patológicas de temor o susto.-

Son respuestas afectivas exageradas a peligros externos reales y actuales. Abarca los reactivos 39 a 46.

6) Síntomas psicósomáticos.-

Son aquellos trastornos en los cuales aparecen sintomatología orgánica y visceral debido a factores emocionales que se manifiestan a través del sistema nervioso autónomo. Cubre los reactivos 47 a 61.

7) Hipocondriasis y Astenia.-

Son los síntomas psicósomáticos que tienen la característica de preocupación e interés obsesivo del sujeto en relación a su salud y a las condiciones de sus órganos. Reactivos del 62 al 68.

8) Síntomas psicósomáticos gastrointestinales.-

Son los trastornos en los cuales se produce sintomatología orgánica que van desde los órganos de la cavidad bucal hasta las expresiones de mal funcionamiento del estómago, intestino delgado y el intestino grueso, así como los principales órganos secretores relacionados con la digestión. Reactivos del 69 al 79.

9) Sensibilidad y desconfianza excesiva.-

Son las reacciones que presentan los sujetos ante a las actitudes o conductas que adoptan las personas a su alrededor y que se caracterizan por considerar dichas actitudes dirigidas hacia sí mismo provocando desconfianza. Reactivos del 80 al 85.

10) Problemas de psicopatía.-

Son patrones repetitivos y persistentes de conducta no agresiva, que viola los derechos de los demás o bien las reglas sociales correspondientes al grupo de la misma edad, manifestándose por violaciones persistentes a las reglas o leyes, vagabundeo, adicción o abuso a sustancias tóxicas, fugas de casa, robo, mentiras serias y constantes fuera y dentro de casa, así como en las relaciones sociales con uno o más camaradas, preocupación por el bienestar de los mismos y sentimientos de culpa cuando se le presiona o hay dificultades. Reactivos del 86 al 101.

3.4 Confiabilidad

Según el manual original (Weider, 1975) esta se obtuvo a través del método de Kuder-Richardson en una aplicación de 1000 sujetos obteniéndose un coeficiente de .95

3.5 Validez

El manual no reporta un estudio de validez de constructo como tal, sin embargo, describe un estudio realizado con 600 sujetos sin disturbios de personalidad, y 400 con disturbios, tomados a partir de 5 diferentes regiones de los Estados Unidos. Al final del capítulo se puede ver la Tabla 1 donde aparece el porcentaje de sujetos normales y los sujetos con desordenes neuropsiquiátricos, quienes fueron clasificados como rechazados neuropsiquiátricos a diferentes niveles de corte (Weider, 1946; Weider, 1975).

Eysenck en una revisión hecha al Cornell Index reporta evidencia escasa de validez concurrente del cuestionario con el MMPI, en donde encontró que este correlaciona con las escalas que forman la triada neurótica y psicótica con un valor de .60, y valores de .30 a .50 en las escalas individuales, sin embargo, en lo referente al manual afirma que este se presenta confuso, no se reporta adecuadamente la confiabilidad y la validez, así como no hay información clara sobre las tablas normativas, cómo fueron obtenidas, y sobre todo cual es el criterio para elaborar puntos de corte tomando en cuenta las contestaciones de los sujetos (Mann, 1950; Buros, 1953).

Según Weider y colaboradores desde que el Cornell Index se utilizó para identificar problemas neuropsiquiátricos, esta tabla se utilizó como evidencia de validez (ver Tabla A).

3.6 Administración

El Cornell Index puede ser autoadministrado o bien aplicado a grupos de cualquier tamaño. Las instrucciones son presentadas de manera clara, y pueden ser leídas por cualquier sujeto que sepa leer y escribir. Al inicio el sujeto deberá llenar datos de tipo demográfico como son su nombre, edad, sexo, ocupación, escolaridad, etc. El cuerpo del Cornell Index consiste en reactivos o preguntas que se refieren a síntomas psiquiátricos y psicósomáticos. No están elaborados

con lenguaje técnico, son reactivos como: *¿Padece mareos frecuentemente?*; *¿Se considera usted nervioso?*.

Las instrucciones para la contestación del cuestionario son las siguientes:

"Trace un círculo alrededor de la palabra SI, cuando su respuesta sea afirmativa. Cuando sea negativa, encierre en el círculo la palabra NO. Conteste todas las preguntas que a continuación se le presentan. Marque la respuesta que le parezca más justa."

El tiempo necesario para contestar el cuestionario varía en función del grado escolar que posee el sujeto al que se le aplica el instrumento, en estudiantes universitarios se reporta de 5 a 7 minutos, en sujetos con menor escolaridad puede llegar a tomar de 10 a 15 minutos (Weider, 1975).

3.7 Calificación

En la versión original se hace a través de plantillas de calificación, en donde se cuenta el número de respuestas afirmativas marcadas por el sujeto en el cuerpo del cuestionario. La calificación toma mas o menos 1 minuto. Después, los valores obtenidos para cada área son transformados a rangos percentilares, según la tabla D elaborada para ello (Weider, 1975) esta tabla se encuentra al final del capítulo.

Según el manual original la calificación puede llevarse a cabo por tres métodos, tomando en cuenta la tabla para niveles de corte elaborada para ello, esta tabla también puede ser revisada al final del capítulo:

Método A: Una calificación en el cuestionario de 23 o más marcas afirmativas sobre la mitad de las personas con serios problemas neuropsiquiátricos y trastornos psicósomáticos, y un número reducido de personas aparentemente sanas.

Método B: Una calificación de 13 o más marcas que la mayoría de las personas con serios problemas neuropsiquiátricos y trastornos psicósomáticos y un número moderado de personas aparentemente sanas.

Método C: Una calificación de 13 o más marcas en los reactivos marcados como "stop questions", que aparecen marcados en las plantillas originales, y que son reactivos considerados como de alarma o indicadores de fuertes problemas neuropsiquiátricos y psicósomáticos (ver Tabla B y C).

3.8 Usos del Cornell Index

El Cornell Index puede ser utilizado tomando en cuenta los reactivos "stop", como una entrevista individual más detallada. Según el autor el instrumento ha mostrado su utilidad para indicar la presencia de estados de ansiedad, hipocondriasis, tendencias antisociales, migraña, asma, problemas gástricos, y síndromes clínicos fronterizos. Es menos efectivo para diagnosticar los llamados "disturbios monosintomáticos", como pueden ser estados prepsicóticos y estados psicóticos iniciales. No es efectivo para diagnosticar estados obsesivos (Wolff, 1946, Weider, 1975).

El instrumento no funciona de manera diferencial como un instrumento para la medición y comparación de diversas características. Desviaciones del promedio, o "normalidad", se reflejan en la calificación del cuestionario, pero esta no refleja las dificultades específicas que padece el sujeto. Por ejemplo sujetos con hipocondriasis, ansiedad, astenia, o personalidades inadecuadas pueden presentar la misma calificación al contestar el cuestionario, la manera de ser más precisos para el diagnóstico diferencial es realizar una entrevista más fina tomando en cuenta las áreas con más problema.

El cuestionario ha sido aplicado en las siguientes áreas:

1. Departamentos de pacientes externos y custodios con problemas psiquiátricos y neurológicos.- Se ha utilizado el instrumento en la clasificación de pacientes externos, y pacientes en custodia que presenta problemas tanto neurológicos como psiquiátricos, el cuestionario se ha utilizado como entrevista estandarizada para el manejo de este tipo de pacientes (Wolff, 1946).

2. Departamentos con pacientes médicos y sometidos a cirugía.- El cuestionario ha sido utilizado con sujetos que presentan diversos problemas emocionales, en los cuales la convalecencia se hace problemática y lenta, el cuestionario fue de gran ayuda para identificar a los pacientes con estos problemas (Weider, 1975).

3. Industria. - Se ha utilizado el cuestionario en los departamentos o secciones médicas de diversas empresas para identificar sujetos con posibles desordenes neuropsiquiátricos y psicósomáticos, en los cuales se presenta baja productividad y poca adaptación al ambiente de trabajo. El cuestionario no se utilizó en el reclutamiento y la selección de personal ya que no era uno de los fines para los que fue creado, y porque no se presentó evidencia de su utilidad en este rubro (Weider, 1975).

4. Investigación. - Se sugiere utilizar el cuestionario con poblaciones mas amplias y en diversas situaciones de tal forma, que se pueda reunir evidencia empírica de su utilidad para la detección de sujetos con problemas de orden neuropsiquiátrico y psicósomático (Mann, 1950; Weider, 1975).

El autor asegura que dada la brevedad del cuestionario, la facilidad para la calificación, y la semejanza en el contenido con una entrevista psiquiátrica, es **adecuado** utilizar el instrumento en diversa situaciones de orden clínico (Weider, 1975).

3.9 Estudios en otros países

El Cornell ha **sido** analizado en otros países, específicamente en España y México, por **ejemplo** este ha sido analizado con sujetos españoles en tres muestras diferentes: Una que constaba de 65 pacientes psiquiátricos; otra con 36 estudiantes y una tercera con 150 sujetos de población general, para dar un total de 251 sujetos. A partir de este estudio se concluyó que el cuestionario es válido para el **diagnóstico** cualitativo de trastornos neuróticos, psicósomáticos y reacciones psicopáticas, esto tomado en cuenta los datos obtenidos a partir de los tres **grupos** anteriores, en donde se observaron diferencias estadísticamente **significativas**, y por los elevados coeficientes de correlación obtenidos en las calificaciones del segundo grupo estudiando, con escalas de calificación clínica de **terapeutas** que entrevistaron a este grupo (Cerdeña, 1957).

Otros estudios llevados con el cuestionario son los realizados en México, en uno de ellos se utilizó el instrumento con 50 sujetos de población marginada de **edades** comprendidas entre 13 y 18 años de edad que tenía por objetivo hacer un estudio exploratorio (Guerrero, 1987). Un segundo estudio fue el realizado por Villegas, Varela y colaboradores en 1993, para obtener datos de confiabilidad a partir de una muestra de 515 sujetos universitarios con edades de 19 a 30 años de

nivel de licenciatura, en donde se concluyó que no todos los reactivos discriminaron y que este debería ser revisado y aplicado con una muestra más amplia (Villegas y Varela, 1993). No se han encontrado más reportes de investigación sobre la confiabilidad y validez del Cornell Index en población mexicana, en la investigación de Varela y Villegas el fin era tratar de ver cuál era la consistencia interna del instrumento.

TABLA A
Porcentaje de aceptados y rechazados psiquiátricos, identificados como rechazados en varios niveles de corte en el Cornell Index.*

Niveles de Corte	400 Rechazados Psiquiátricos	600 Aceptados Psiquiátricos
0	100 %	100 %
1	99	82
2	97	67
3	94	54
4	93	46
5	92	39
6	90	32
7	86	28
8	85	24
9	83	20
10	81	18
11	78	16
12	76	15
13	74	13
14	72	12
15	68	10
16	66	9
17	62	8
18	61	7
19	60	7
20	57	6
21	55	5
22	53	4
23	50	4
24	48	4
25	45	3
26	42	3
27	41	3
28	40	2
29	39	1
30	35	1
31	34	1
32	32	1

*En términos de la opinión en una entrevista psiquiátrica durante 5 sesiones.

La tabla se lee como sigue: Si se usa una calificación de 7 en el *Index*, 86% de los rechazados en la entrevista podrían ser rechazados en el *Index*; 28% de los aceptados en la entrevista podrían también ser rechazados por el *Index*.

TABLA B

Porcentaje de aceptados y rechazados psiquiátricos identificados como rechazados en varios niveles de corte en el Cornell Index, usando la calificación total de uno más reactivos "stop".*

Niveles de Corte, o uno o más reactivos "stop"	400 Rechazados Psiquiátricos	600 Aceptados Psiquiátricos
0	100 %	100 %
1	99	81
2	97	67
3	95	55
4	94	48
5	93	42
6	91	36
7	90	32
8	90	28
9	88	25
10	86	24
11	85	22
12	84	21
13	83	20
14	82	19
15	80	18
16	78	18
17	76	17
18	75	16
19	75	16
20	74	16
21	73	15
22	72	15
23	71	14
24	69	14
25	69	14
26	68	14
27	68	14
28	67	14
29	67	14
30	66	14
31	64	14
32	64	14

*En términos de la opinión en una entrevista psiquiátrica durante 5 sesiones.

La tabla se lee como sigue: si en una calificación de corte de siete en el *Index* o uno o más reactivos "stop" a sido usado, 90% de los rechazados en la entrevista podrian ser rechazados por el *Index*; 32% de los aceptados en la entrevista podrian ser rechazados por el *Index*.

TABLA C

*Porcentaje de aceptados y rechazados psiquiátricos * identificados como rechazados por los 3 métodos de calificación del Cornell Index.*

Método de Calificación	400 Rechazados Psiquiátricos	600 Aceptados Psiquiátricos
Método A 23 o más reactivos	50%	4%
Método B 13 o más reactivos	74	13
Método C 13 o más reactivos, uno o más reactivos "stop"	83	20

*En términos de la opinión en una entrevista psiquiátrica durante 5 sesiones.

TABLA D

Normas percentilares para 1,298 universitarios de primer ingreso, no seleccionados con recomendación para problemas neuropsiquiátricos

Percentil	Calificación	
	Hombres	Mujeres
99	0	0
95	0	0
90	1	1
80	2	2
75	2	3
70	3	4
60	4	5
50	6	6
40	7	8
30	9	10
25	11	12
20	12	13
10	17	17
5	21	22
1	29	34
No. de casos	836	462
Media	7.35	8.42
Desviación Estandard	6.76	7.48

La distribución semejante para hombres y mujeres puede llegar a constituir una justa evidencia en de que las diferencias de sexo en la calificación total sobre la prueba son insignificantes, aunque la permanece la posibilidad de que la significación diagnóstica de una calificación dada puede diferir de acuerdo al sexo del examinado.

Nota: Las 4 Tablas anteriores (A, B, C, D) fueron tomadas del Manual original del Cornell Index (Weider, 1975).

Capítulo IV

METODO

4.1 Preguntas de investigación.

- ¿La versión utilizada del Cornell Index es entendida por los estudiantes de Ciudad Universitaria de ambos sexos?
- ¿Cuales son las características de Confiabilidad y Validez del Cornell Index al ser aplicado en una muestra de estudiantes universitarios?

4.2 Hipótesis de investigación.

- Los reactivos que forman la versión en español del Cornell Index, son entendidos en un 100% por la muestra de estudiantes universitarios empleada.
- Existen diferencias entre los datos de Confiabilidad y Validez de la versión original y los datos de Confiabilidad y Validez obtenidos a partir de la aplicación a la muestra de estudiantes universitarios mexicanos.

4.3 Variables.

- Variable Independiente: Es la muestra de estudiantes universitarios, de ambos sexos, extraídos de las diferentes facultades y escuelas de la U.N.A.M.
- Variable Dependiente: Las escalas en la aplicación del Cornell Index a la muestra de estudiantes universitarios mexicanos.

4.4 Población.

- Esta formada por la población de estudiantes de Ciudad Universitaria, inscritos en el año de 1995.

4.5 Muestra

- La muestra fue proporcional, seleccionada al azar de cada una de las 14 escuelas y facultades que conforman la Ciudad Universitaria se obtuvo a partir de la agenda estadística del año de 1995 que edita la U.N.A.M. La muestra estuvo compuesta por un total de 765 sujetos, todos ellos estudiantes de la licenciatura con edades comprendidas de entre 18 y 25 años.

4.6 Escenario

- El lugar donde se realizaron las aplicaciones de los cuestionarios fueron los salones de clase de las distintas escuelas y facultades en la Ciudad Universitaria, siempre con la ayuda del profesor en turno.

4.7 Instrumento

Para la aplicación se eligió el instrumento llamado Cornell Index en la versión que se utiliza en las prácticas de Evaluación de la Personalidad en el sexto semestre de la licenciatura en Psicología. Dicha versión fue editada por Angelina Guerrero en el año de 1990 (ver Apéndice).

Descripción del Instrumento: El Cornell Index es un instrumento autodescriptivo de la personalidad, que puede ser administrado a adolescentes y adultos. Consta de 101 reactivos que examinan de manera global aspectos de personalidad que se presentan como síntomas psicossomáticos y neuropsiquiátricos, tiene la finalidad de funcionar como historia psiquiátrica estandarizada, y como una guía de entrevista clínica en general, tratando de establecer diferencias entre personas con padecimientos psicossomáticos y neuropsiquiátricos de las personas con problemas personales de la población en general. Fue creado entre 1944 y 1949 por Weider y colaboradores a partir del Cornell Select Index, forma N, para uso civil aplicable a hombres y mujeres de entre 18 y 40 años de edad, y el Cornell Service Index que incluye alrededor de 30 reactivos en relación directa con el comportamiento del individuo en el servicio militar. Al finalizar la Segunda Guerra Mundial los instrumentos para el servicio militar fueron modificados para su uso en población civil. Los reactivos del Cornell Service Index fueron validados en población civil, los que tenían mayor significancia para uso clínico fueron combinados con los reactivos del Cornell Select Index para así lograr un Cornell Index más eficaz esta última forma es la que se conoce como Cornell Index FN2, que consta de 64 reactivos de la forma original N, más 36 adicionales y 1 introductorio, dando un total de 101 reactivos (Buros, 1953, 1959, 1974, 1978; Cerda 1957, 1977; Mittelman, 1946).

El Cornell Index evalúa los siguientes aspectos:

- ◆ Problemas de adaptación debidos a sentimientos de miedo e insuficiencia.
- ◆ Reacciones patológicas del humor especialmente la depresión.
- ◆ Angustia.

- ◆ Síntomas psicósomáticos neurocirculatorios.
- ◆ Reacciones patológicas de temor.
- ◆ Síntomas psicósomáticos.
- ◆ Hipocondría y astenia.
- ◆ Síntomas psicósomáticos gastrointestinales.
- ◆ Sensibilidad y desconfianza excesiva.
- ◆ Problemas psicopáticos y sociopáticos.

Estas áreas poseen clínicamente tres grados de severidad: Leve, Moderado y Severo, esto de acuerdo al manual del instrumento en la versión utilizada para esta investigación. En dicho manual no aparecen datos de Confiabilidad y Validez, tampoco aparecen datos sobre los grados de severidad para el diagnóstico; únicamente se utilizó el cuestionario que se presenta en dicho material, para este cuestionario se tuvo que diseñar una hoja de respuestas para que el sujeto pudiese contestar aparte, y el manejo de la información fuese más adecuado.

La confiabilidad del cuestionario original es reportada por el autor como de .95 obtenida a partir del método de Kuder-Richardson. En lo referente a la validez del Cornell Index este fue correlacionado con el MMPI, obteniéndose datos tanto para triada neurótica como psicótica de .60, y en las escalas individuales valores de .30 a .50. (Mittelman, 1946; Weider, 1946; Mann, 1950; Weider, 1975).

4.8 Diseño

- Es un diseño de una muestra con una sola medición (aplicación de la forma única que compone el material)

4.9 Tipo de Estudio

- El tipo de estudio fue exposfacto evaluativo, descriptivo de campo, ya que no se realizó manipulación de variables, solo existe la asignación de los sujetos al azar. Descriptivo, porque se describirán las características que presentaron los estudiantes al contestar el cuestionario. De campo, ya que se hizo la aplicación en escenarios naturales.

4.10 Procedimiento

Se procedió a obtener la muestra a partir de la población de estudiantes de Ciudad Universitaria, tomando como base los datos que aparecen en la agenda estadística 1995, que edita la misma Universidad. La muestra debería de tener como característica que los estudiantes pertenecieran al nivel de la licenciatura, fuesen de sexo masculino o femenino y que tuviesen edades de entre 18 y 25 años.

Una vez obtenida la muestra se giraron oficios a los respectivos secretarios generales de cada escuela y facultad, con el fin de obtener su consentimiento para la aplicación del cuestionario. La aplicación se hizo de manera colectiva, proporcionando todo el material necesario (cuadernillos, hojas de respuesta y lápices). Se administró a cada sujeto el instrumento completo en una sola sesión. Durante la administración se tomó en cuenta el tiempo que tardaban los estudiantes en contestar, así como los reactivos que se presentaron como confusos o poco entendibles al ser contestados. Cabe aclarar que el número total de la muestra no fue el que originalmente había sido calculado, ya que se presentaron problemas para la aplicación al momento de presentarse a las escuelas, en algunas de ellas no fue posible administrar el cuestionario, debido a los contratiempos administrativos ó simplemente de negación a que fuese aplicado el instrumento.

Capítulo V

**RESULTADOS, DISCUSION
Y
CONCLUSIONES**

5.1 Análisis Estadístico

Una vez administradas y calificadas las pruebas, se procedió a la codificación de los datos, para posteriormente realizar los respectivos análisis estadísticos. Como regla general se procedió a realizar un análisis de frecuencias, esto con la finalidad de saber como se distribuían las calificaciones de los sujetos, además de saber cuáles eran los datos demográficos de la muestra de estudio. Se realizó el Análisis de Consistencia Interna para saber el grado de exactitud del instrumento. Por último, se hizo el cálculo de la Validez con el objeto de comprobar si el instrumento mide lo que dice medir (Validez de Constructo).

Para el análisis estadístico se utilizó el programa de computo SPSS versión 5.1 para Windows, específicamente para el análisis de frecuencias y el cálculo de la Confiabilidad, en el caso de la Validez se utilizó el programa STATISTICA versión 4 para Windows.

5.2 Descripción de la muestra

La muestra estudiada estuvo compuesta por 765 sujetos, con edades de entre 18 y 25 años de edad, 436 hombres y 329 mujeres, todos ellos alumnos de las distintas licenciaturas que se imparten en Ciudad Universitaria. Cabe aclarar que el número que formó la muestra de investigación tiene por objetivo hacer que los datos se distribuyan lo mejor posible, según Nunnally (1995) lo más conveniente en el análisis de un instrumento es el aplicar 10 pruebas por cada reactivo, sin embargo, es posible aplicar 5 pruebas por cada reactivo, en el caso específico de esta investigación se realizaron 765 aplicaciones, cumpliendo cabalmente con lo anterior.

5.3 Poder discriminativo y Confiabilidad

Al llevar a cabo el análisis de reactivos, que se refiere a si los reactivos que componen el instrumento son sensibles, se tomó como criterio de exclusión el valor de .20 o más para tomar en cuenta si el reactivo discrimina o no, aunque cabe aclarar que el ideal es que los reactivos presenten valores superiores a .30, los que presentan valores de .20 serán considerados como moderadamente discriminativos. Estos valores fueron obtenidos por la correlación biserial puntual para el método ítem total. En el caso de la Confiabilidad, Nunnally (1995) afirma que para que un instrumento se considere como confiable es necesario que este cuente con valores arriba de .85 de Alpha, sin embargo, Magnusson (1975) apoya

la tesis de que si la varianza verdadera en un instrumento es mayor a .51, y menor a este valor la varianza de error, ya se puede considerar dicho instrumento como confiable, con sus respectivas reservas. A continuación se presentaran una serie de tablas que muestran los diversos valores de la relación ítem-total y confiabilidad que resultaron del análisis de las distintas áreas que abarca el Cornell Index. El análisis de reactivos se hizo a través de la correlación ítem-total, y la consistencia interna de cada área que mide el instrumento se hizo a través del Alpha de Cronbach (Nunnally, 1995; Aiken, 1995; Reidl, 1990):

Area I (DA) Deficiencia de adaptación que se experimenta por sentimientos de miedo e insuficiencia.- Para esta área sólo discriminaron 13 reactivos de 18 que la componen, con un Alpha de .6733.

TABLA 1

Area I (DA)
N= 765
Media= 3.4496
Desv. Estándar 2.6657
Alpha= .6733

Reactivo No.	Media	Desv. Estándar	Correlación Item-Total
R2	.2195	.4143	.3133*
R3	.2407	.4279	.2315*
R4	.0726	.2597	.1186
R5	.1540	.3613	.2906*
R6	.0973	.2967	.2401*
R7	.2319	.4224	.3950*
R8	.1044	.3061	.1167
R9	.4407	.4969	.3459*
R10	.1487	.3561	.2288*
R11	.0584	.2347	.1853
R12	.0690	.2537	.2624*
R13	.2584	.4381	.3544*
R14	.2425	.4290	.2737*
R15	.3097	.4628	.3139*
R16	.3912	.4884	.2850*
R17	.1681	.3743	.1396
R18	.0991	.2991	.1790
R19	.1434	.3508	.3722*

*Los reactivos marcados con asterisco, son los que discriminaron adecuadamente.

Area II (D) Reacciones patológicas del humor, en especial la Depresión.- El área de compone de 7 reactivos, de los cuales sólo discriminaron 5, con un valor Alpha de .4781

TABLA 2

Area II (D)
N= 765
Media= 1.4035
Desv. Estándar 1.3382
Alpha= .4781

Reactivo No.	Media	Desv. Estándar	Correlación Item-Total
R20	.1681	.3743	.3270*
R21	.0690	.2537	.3043*
R22	.1628	.3695	.1926
R23	.1434	.3508	.2243*
R24	.2531	.4352	.1121
R25	.2035	.4030	.2338*
R26	.4035	.4910	.2567*

*Los reactivos marcados con asterisco, son los que discriminaron adecuadamente.

Area III (A) Angustia.- En esta área, que se compone de 7 reactivos, sólo discriminaron 6, con un valor de Alpha de .5759.

TABLA 3

Area III (A)
N= 765
Media= 1.6531
Desv. Estándar 1.5437
Alpha= 5759

Reactivo No.	Media	Desv. Estándar	Correlación Item-Total
R27	.4496	.4479	.3556*
R28	.0850	.2791	.2211*
R29	.1434	.3508	.2109*
R30	.3062	.4613	.3989*
R31	.2690	.4438	.4431*
R32	.1363	.3434	.2592*
R33	.0230	.1501	.1705

*Los reactivos marcados con asterisco, son los que discriminaron adecuadamente.

Area IV (PN) Síntomas Psicósomáticos Neurocirculatorios.- En esta área, compuesta por 5 reactivos, discriminaron todos, hay un valor del Alpha de .5966.

TABLA 4

Area IV (PN)
 N= 765
 Media= .4814
 Desv. Estándar .9120
 Alpha= .5966

Reactivo No.	Media	Desv. Estándar	Correlación Item-Total
R34	.0903	.2868	.2827*
R35	.0903	.2868	.4186*
R36	.0867	.2817	.3282*
R37	.1257	.3318	.3600*
R38	.0885	.2843	.3775*

*Los reactivos marcados con asterisco, son los que discriminaron adecuadamente.

Area V (T) Reacciones Patológicas de Temor.- El área se compone de 8 reactivos, sólo discriminaron 6, presenta un valor de Alpha de .5161.

TABLA 5

Area V (T)
 N= 765
 Media= 1.8726
 Desv. Estándar 1.5442
 Alpha= .5161

Reactivo No.	Media	Desv. Estándar	Correlación Item-Total
R39	.1876	.3907	.3326*
R40	.0920	.2893	.1514
R41	.0923	.2868	.1557
R42	.2779	.4483	.2284*
R43	.5168	.5002	.2004*
R44	.1841	.3879	.2597*
R45	.2619	.4401	.3257*
R46	.2619	.4401	.2607*

*Los reactivos marcados con asterisco, son los que discriminaron adecuadamente.

Area VI (P) Síntomas Psicossomáticos.- El área se compone de 15 reactivos, sólo discriminaron 5, presentó un Alpha de .4523.

TABLA 6

Area VI (P)
N= 765
Media= 1.4177
Desv. Estándar 1.4632
Alpha= .4523

Reactivo No.	Media	Desv. Estándar	Correlación Item-Total
R47	.1805	.3850	.2359*
R48	.1150	.3194	.2567*
R49	.0513	.2209	.2025*
R50	.0938	.2918	.1298
R51	.2106	.4081	.1064
R52	.0779	.2682	.0419
R53	.1681	.3743	.1185
R54	.0230	.1501	.1389
R55	.0088	.0937	.0254
R56	.0336	.1804	.0251
R57	.0372	.1893	.0390
R58	.1027	.3038	.1988
R59	.0938	.2918	.2820
R60	.1097	.3128	.2377
R61	.1115	.3150	.1748

*Los reactivos marcados con asterisco, son los que discriminaron adecuadamente.

Area VII (HA) Hipocondría y Astenia.- Esta área se compone de 7 reactivos, discriminaron solamente 2, el valor de Alpha es de .3770

TABLA 7

Area VII (HA)
N= 765
Media= 1.1717
Desv. Estándar 1.1348
Alpha= .3770

Reactivo No.	Media	Desv. Estándar	Correlación Ítem-Total
R62	.3805	.4859	.1947
R63	.0779	.2682	.1101
R64	.0779	.2682	.1037
R65	.2230	.4166	.2664*
R66	.2956	.4567	.1764
R67	.0761	.2654	.1630
R68	.0407	.1978	.2080*

*Los reactivos marcados con asterisco, son los que discriminaron adecuadamente.

Area VIII (PG) Síntomas Psicósomáticos Gastrointestinales.- Esta compuesta por 11 reactivos, de los cuales discriminaron 10, con un valor de Alpha de .6768

TABLA 8

Area VIII (PG)
 N= 765
 Media= .8920
 Desv. Estándar 1.4516
 Alpha= .6768

Reactivo No.	Media	Desv. Estándar	Correlación Item-Total
R69	.1009	.3014	.2164*
R70	.1239	.3298	.3790*
R71	.1027	.3038	.3416*
R72	.0531	.2244	.2274*
R73	.0761	.2654	.4087*
R74	.0549	.2279	.3930*
R75	.1150	.3194	.5027*
R76	.0283	.1660	.2766*
R77	.0584	.2347	.3371*
R78	.1186	.3236	.3397*
R79	.0602	.2380	.1697

*Los reactivos marcados con asterisco, son los que discriminaron adecuadamente.

Area IX (SD) Sensibilidad y Desconfianza Excesiva.- El área se compone de 8 reactivos, discriminaron únicamente 3, se obtuvo un Alpha de .4399

TABLA 9

Area IX (SD)
 N= 765
 Media= 1.2319
 Desv. Estándar 1.2003
 Alpha= .4399

Reactivo No.	Media	Desv. Estándar	Correlación Item-Total
R80	.1646	.3711	.2533*
R81	.1080	.3106	.2954*
R82	.0885	.2843	.1661
R83	.2549	.4362	.1890
R84	.2088	.4068	.2319*
R85	.4071	.4917	.1912

*Los reactivos marcados con asterisco, son los que discriminaron adecuadamente.

Area X (Ppt) Problemas Psicopáticos y Sociopáticos. - El área se compone de 16 reactivos, solo discriminaron 11, se obtuvo un Alpha de .6660

TABLA 10

Area X (PPT)
N= 765
Media= 3.0106
Desv. Estándar 2.4108
Alpha= .6660

Reactivo No.	Media	Desv. Estándar	Correlación Item-Total
R86	.3310	.4710	.3297*
R87	.1947	.3963	.2009*
R88	.1770	.3820	.4273*
R89	.0265	.1609	.1305
R90	.0230	.1501	.1345
R91	.0319	.1758	.1031
R92	.0230	.1501	.1796
R93	.0850	.2791	.2137*
R94	.4053	.4914	.2537*
R95	.0885	.2843	.1671
R96	.3487	.4770	.3450*
R97	.1646	.3711	.3658*
R98	.2619	.4401	.3660*
R99	.0655	.2476	.2138*
R100	.1009	.3014	.3098*
R101	.2425	.4290	.4146*

*Los reactivos marcados con asterisco, son los que discriminaron adecuadamente.

5.4 Validez

El cálculo de la Validez se hizo a través de una Análisis Factorial, de componentes principales, este procedimiento tiene como fin el tratar de ver si los reactivos que componen el instrumento se agrupan como originalmente se planteó en su construcción, o bien, saber si estructuralmente lo hacen de manera distinta a como se creó la prueba. Los resultados de esta investigación mostraron 36 Factores con valores Eigen superiores a 1.0, sin embargo, sólo se tomarán en cuenta los 10 primeros factores, ya que son los que explican mas ampliamente la varianza total del instrumento, los restantes 26 Factores aunque tienen valores Eigen superiores a 1, no cargaron como los 10 primeros, y se consideró que eran factores residuales. A continuación se presenta la Tabla 11 y 12, en donde se muestran los valores Eigen para estos 10 Factores, así como las cargas factoriales en los 10 primeros factores, se tomarán en cuenta cargas factoriales superiores a .39, este criterio, aunque resulta ser arbitrario, autores abajo citados dan apoyo a esta decisión. Nunnally (1995) y Martínez (1995) plantean que la opinión más frecuente para realizar un Análisis Factorial es que sean 5 sujetos por variable. Guadanoli y Velicer (1988, citados por Martínez, 1995) basándose en datos de simulación, señalan que para que los factores puedan considerarse confiables, la decisión no sólo debe basarse en el tamaño muestral, sino que deben examinarse las saturaciones en los factores. Proponen las siguientes reglas:

- 1) Los Factores (componentes) con cuatro o mas saturaciones por encima de .60 pueden considerarse confiables, independientemente del tamaño de la muestra.
- 2) Factores con 10 o mas saturaciones bajas (alrededor de .40) son confiables siempre que el tamaño de la muestra sea superior a 150.
- 3) Factores con saturaciones factoriales bajas, no serán interpretados a menos que el tamaño de la muestra sea de 300.

TABLA 11
ANALISIS FACTORIAL
VALORES EIGEN
EXTRACCION: COMPONENTES PRINCIPALES

VALOR	VALOR EIGEN	VALORES EIGEN ACUMULADOS	VARIANZA ACUMULADA
1	9.245391	9.24539	9.24539
2	3.668873	12.91426	12.91426
3	2.650291	15.56456	15.56456
4	2.387997	17.95255	17.95255
5	2.276348	20.22890	20.22890
6	2.013509	22.24241	22.24241
7	1.916504	24.15891	24.15891
8	1.876584	26.03550	26.03550
9	1.821189	27.85669	27.85669
10	1.738487	29.59517	29.59517

TABLA 12
Resultados del Análisis Factorial para la validez del Cornell Index

VARIABLE	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6	FACTOR 7	FACTOR 8	FACTOR 9	FACTOR 10
R2						38563				
R3										
R4										
R5										32808
R6					54721					
R7			30476							47957
R8								31359		
R9										42791
R10										
R11							30434	34684		
R12		39667								
R13										56557
R14										31797
R15										42857
R16										37509
R17			30921							
R18										
R19										
R20						42239				
R21						37964				
R22										
R23								32041		
R24										
R25			52837							
R26										
R27									31907	
R28									40958	
R29										
R30									38241	30519
R31									35149	
R32									44368	
R33					33049				46428	
R34										
R35					53899					
R36					49708					

TABLA 12
Resultados del Análisis Factorial para la validez del Cornell Index

VARIABLE	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6	FACTOR 7	FACTOR 8	FACTOR 9	FACTOR 10
R37					49426					
R38					56771					
R39										
R40										
R41										
R42										47765
R43		-32100								33270
R44										
R45	33339					37814				
R46	35977									
R47				34166	31348					
R48										
R49		35378								
R50										
R51										
R52										
R53										
R54							46574			
R55									32837	
R56										
R57										
R58						33524				
R59						33024				
R60	41675									
R61		36674							38687	
R62	34208									
R63										
R64						32633				
R65										
R66	32089									
R67						43934				
R68						59203				
R69						59635				
R70				45276						
R71				39592						

TABLA 12
Resultados del Análisis Factorial para la validez del Cornell Index

VARIABLE	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6	FACTOR 7	FACTOR 8	FACTOR 9	FACTOR 10
R72									- 31811	
R73				64887						
R74				31784				52352		
R75				69264						
R76								60544		
R77				47283						
R78				58830						
R79								42189		
R80	35826									
R81										
R82		35397								
R83										37634
R84										
R85				51370						
R86	47030									
R87				58467						
R88	61430									
R89								54934		
R90								55810		
R91								35302		
R92		49359								
R93		51336								
R94	48273									
R95		57160								
R96	42707									
R97		35253		31658						
R98	52525									
R99		31636								
R100		31234								
R101	58026									

5.5 Análisis de Resultados y Discusión.

5.5.1. Confiabilidad

A partir de los análisis realizados en lo referente a la obtención del poder discriminativo de cada uno de los reactivos que conforman las diferentes áreas del Cornell Index y la confiabilidad, a través de la consistencia interna por el método de Cronbach se encontró lo siguiente:

Area I (DA) Deficiencia en la adaptación que se experimenta por sentimientos de miedo e insuficiencia.- Esta área se compone de 18 reactivos, de los cuales sólo discriminaron 13: R2, R3, R5, R6, R7, R9, R10, R12, R13, R14, R15, R16 y R19; con un valor de Alpha de Cronbach=.6733.

Area II (D) Reacciones patológicas del humor, especialmente la depresión.- El área se compone de 7 reactivos, sólo discriminaron 5: R20, R21, R23, R25 y R26; con un valor de Alpha de Cronbach =.4781, la confiabilidad para esta área es baja, por tanto los reactivos que la componen no son para uso adecuado.

Area III (A) Angustia.- Esta área se compone de 7 reactivos, de los cuales solo discriminaron 6: R27, R28, R29, R30, R31 y R32; la confiabilidad presento un Alpha=.5759, esta área también posee moderada confiabilidad.

Area IV (PN) Síntomas psicósomáticos neuroculatorios.- El área se compone de 5 reactivos, en esta discriminaron todos: R34, R35, R36, R37 y R38; el valor de Alpha para esta escala es de .5966, posee moderada confiabilidad.

Area V (T) Reacciones patológicas de temor.- El área se compone de 8 reactivos, de los cuales solo discriminaron 6: R39, R42, R43, R44, R45, y R46; en la confiabilidad arrojó un Alpha=.51, escasamente resulta confiable.

Area VI (P) Síntomas psicósomáticos.- Esta área se compone de 15 reactivos, de los cuales solo discriminaron 5: R47, R48, R49, R59 y R60; presento un Alpha=.4523, la confiabilidad para esta área no es adecuada.

Area VII (HA) Hipocondría y Astenia.- 7 reactivos conforman esta área de los cuales solo discriminaron 2: R65 y R68; tampoco la confiabilidad fue adecuada, arrojó un Alpha=.3770.

Area VIII (PG) Síntomas psicósomáticos gastrointestinales.- El área esta compuesta por 11 reactivos discriminaron 10: R69, R70, R71, R72, R73, R74, R75, R76, R77 y R78; la confiabilidad presento un valor de $\text{Alpha}=.6768$, es moderadamente confiable.

Area IX (SD) Sensibilidad y desconfianza excesiva.- El área esta formada por 6 reactivos, de los cuales solo 3 discriminaron: R80, R81 y R84; el valor de Alpha para esta área es de .4399, no posee confiabilidad adecuada.

Area X (PPI) Problemas psicopáticos y sociopáticos.- El área la conforman 16 reactivos, discriminaron 11: R86, R87, R88, R93, R94, R96, R97, R98, R99, R100 y R101; se obtuvo un $\text{Alpha}=.6660$, la confiabilidad para el área es moderada.

Los datos anteriores muestran que no todos los reactivos que forman la versión del Cornell Index utilizada para esta investigación poseen poder discriminativo, así mismo, ninguna de las diferentes áreas que supuestamente mide el instrumento son adecuadamente confiables, si se toma en cuenta el criterio de que una confiabilidad significativa para un instrumento debe ser de .85 (Nunnally, 1995). Sin embargo, se puede afirmar que las 6 áreas que presentaron valores $\text{Alpha} \geq .51$ poseen confiabilidad muy baja (Magnusson, 1973) por ello no se puede hablar de valores consistentes.

5.5.2. Validez

A partir del análisis factorial con extracción de componentes principales y rotación varimax llevado a cabo para este instrumento, se obtuvieron 36 factores con valores Eigen superiores a 1, pero sólo se tomaron en cuenta los primeros 10 factores, ya que fueron los que presentaron mayores cargas factoriales. En total cargaron 80 reactivos de los 100 que fueron analizados, los 10 principales factores quedaron conformados de la siguiente manera:

- Factor 1.- Quedo formado por los reactivos: 45, 46, 60, 62, 66, 80, 86, 88, 94, 96, 98, 101.
- Factor 2.- Lo forman los reactivos: 12, 43, 49, 61, 82, 92, 93, 95, 97, 99 y 100.
- Factor 3.- Lo forman los reactivos: 7, 17, 25, 85, 87, 97.
- Factor 4.- Lo forman los reactivos: 47, 70, 71, 73, 74, 75, 77 y 78.

- Factor 5.- Lo forman los reactivos: 6, 33, 35, 36, 37, 38, 47.
- Factor 6.- Lo forman los reactivos: 2, 20, 21, 45, 58, 59, 64, 67, 68, 69.
- Factor 7.- Quedo formado por los reactivos: 11, 54, 89, 90 y 91.
- Factor 8.- Lo forman los reactivos: 8, 11, 23, 61, 74, 76, 79.
- Factor 9.- Esta conformado por los reactivos: 27, 28, 30,31, 32, 33, 55, 72.
- Factor 10.- Lo forman los reactivos: 5, 7, 9, 13, 14, 15, 16, 30, 42, 43, 83.

Los factores arriba descritos muestran que la estructura factorial del Cornell Index no concuerda con la estructura original de la prueba, ya que pareciera ser que se formaron 10 nuevas áreas, esto se debe muy posiblemente a que la versión utilizada para esta investigación desde su origen no cumplió con los requisitos mínimos de adaptación a una nueva cultura (Butcher y Panchieri, 1976).

Dentro de los objetivos de esta investigación está el comprobar en primer lugar si los reactivos que conforman la versión del Cornell Index administrada a la muestra de estudiantes universitarios eran entendidos en su totalidad, se pudo comprobar durante la administración del instrumento que los reactivos sí fueron entendidos, ninguno de los sujetos que participaron en esta investigación reportó que no se entendiesen estos.

Como **segundo** objetivo se necesitaba comprobar si había diferencias entre los datos de confiabilidad y validez de la versión original, con los datos de confiabilidad y validez de la versión utilizada para esta investigación. Al haber hecho una revisión exhaustiva de la literatura existente sobre el Cornell Index se comprobó, por un lado, que la versión original carecía de datos suficientes sobre la confiabilidad y validez, dado que es un instrumento creado en la década de los 40's, este cumplió sus objetivos de creación para esa época, sin embargo, hoy en día tendría que ser revisado para que pudiese ser utilizado en la época actual.

En relación a la versión utilizada en nuestro país, también se hizo una investigación exhaustiva, no encontrando datos que apoyasen el uso adecuado para la cultura mexicana, de ahí la necesidad de obtener datos que nos permitiesen saber si el Cornell Index puede ser utilizado con nuestra población.

Los datos obtenidos mostraron que 80 de los 100 reactivos que conforman el instrumento son sensibles para diferenciar entre sujetos normales y sujetos con problemas neuropsiquiátricos y psicósomáticos. Pero en lo referente a la confiabilidad el instrumento no es adecuado, ya que esta presentó coeficientes muy bajos o deficientes.

Sobre la validez los datos mostraron que la estructura factorial es diferente comparada con la versión original, aunque la versión original no fue sometida a un análisis factorial para comprobar la validez de constructo, sólo muestra validez concurrente; en la versión utilizada para México se encontraron 10 posibles áreas nuevas que darían por resultado un instrumento diferente al Cornell Index. Por lo tanto, se puede afirmar que la versión utilizada para nuestra cultura posee una estructura factorial muy diferente a la que supuestamente tiene la versión original.

5.6 Conclusiones

Tomando en cuenta todo lo anterior, se concluye lo siguiente:

- ◆ De los reactivos que conforman el instrumento llamado Cornell Index, específicamente la forma FN2 sólo 80 de los 100 fueron sensibles para discriminar sujetos normales de aquellos que tienen problemas neuropsiquiátricos y psicósomáticos.
- ◆ La confiabilidad de las distintas áreas que evalúa el Cornell Index se presentó como deficiente o muy baja.
- ◆ La estructura factorial de la versión utilizada para esta investigación es distinta a la versión original, se puede señalar que los datos obtenidos muestran un instrumento con áreas diferentes de evaluación, a la versión original.
- ◆ El manual original del instrumento investigado presenta una tabla normativa expresada en Rangos Percentilares obtenida a partir de una muestra de estudiantes universitarios norteamericanos; en el caso del manual de la versión utilizada para esta investigación este no cuenta con una tabla normativa adecuada, así, la interpretación correcta de los datos no es posible. Por lo tanto, un diagnóstico hecho a partir de esta versión es incorrecto.
- ◆ El llamado Cornell Index, forma FN2, no debe ser utilizado para evaluación en sujetos mexicanos, dado que presenta problemas de adaptación y estandarización.

que no fueron cubiertos desde un principio. De ser utilizada esta versión se podría incurrir en errores graves de diagnóstico, con las consecuencias inherentes a esto.

La idea original del autor del Cornell Index (Weider, 1975) no debe ser desechada, por el contrario nos proporciona un ejemplo magnífico que desde la década de los años 40 ya se estaba haciendo investigación en lo que hoy puede llamarse Psicología de la Salud. Se trataba de contar con instrumentos que pudieran proporcionar de una manera rápida y eficiente características de personalidad que permitiesen diferenciar sujetos con problemas neuropsiquiátricos y psicósomáticos de aquellos que eran normales y se adaptaban mejor al medio. Hoy en día este tipo de instrumentos proporcionaría datos importantes sobre como algunas características de personalidad influyen para que se den ciertas enfermedades, ayudan a poder detectar qué sujetos están más predispuestos a enfermarse que otros, o bien, ayudaría en el manejo de desordenes específicos como fobias, dolor severo, o síndromes de estrés postraumático (Holtzman, 1989).

El estudio de la personalidad hoy por hoy, sigue siendo uno de los campos dentro de la Psicología más fascinantes, pero que más polémica levanta cuando se habla de ella, definirla, clasificarla y medirla, no ha sido una tarea fácil, a pesar de los avances tecnológicos, como es el uso tan amplio en nuestros días de la computadora, el campo de estudio de la personalidad no está agotado, por el contrario, pareciese ser que a más investigaciones, más preguntas aparecen que obligan a corregir y ampliar lo que de cierta forma ya se daba por establecido.

Para poder hacer estudios adecuados en el campo de la personalidad, es imprescindible que contemos con herramientas adecuadas, estas herramientas son los llamados tests psicológicos.

En nuestro país no se cuenta con un monitoreo adecuado para el uso de los tests, investigaciones referentes a esto señalan que el 95% de los instrumentos que se manejan a nivel nacional son de origen extranjero, para los cuales se tiene muy pocos datos de confiabilidad y validez que apoyen su uso en la cultura mexicana. Sólo el 5% de estos fueron desarrollados por mexicanos para mexicanos (Hernández, 1995).

Díaz-Guerrero (1971) afirmaba desde hace mucho: *"el psicólogo en desarrollo debe pues, dedicar su atención a su propia cultura, a semejanza de las construcciones empíricas desarrolladas en la cultura angloamericana deberá escudriñar las propias maneras de ser de su propia gente y desarrollar conceptos que les vengan a sus naturalezas específicas e idiosincráticas"*, por otro lado

afirma "el psicólogo en desarrollo deberá pensar en como puede construir pruebas que valgan total y específicamente para la idiosincrasia de su propio pueblo".

Idealmente los psicólogos en nuestro país deben de construir sus propios instrumentos, pero problemas diversos que van desde la ignorancia en este campo hasta los económicos han impedido que esto se lleve a cabo.

Si es necesario utilizar instrumentos de origen extranjero es indispensable hacer investigaciones que nos permitan comprobar si estos son adecuados para nuestra cultura o no. Estos estudios deben de contemplar la obtención de datos sobre la contabilidad y la validez al ser aplicados a poblaciones distintas para las que originalmente fueron creados. El uso de instrumentos sin haber sido analizados con el debido rigor metodológico trae consecuencias catastróficas, sobre todo para aquellas personas que son evaluadas, sin saber si la herramienta que estamos utilizando sirve para nuestros propósitos o no.

Diariamente se toman decisiones a partir de la aplicación de tests psicológicos, se decide el destino de mucha gente, si esto lo hacemos sin tomar en cuenta que las herramientas cumplan con requisitos mínimos de confiabilidad y validez, incurrimos en problemas éticos muy serios, decidimos la vida de la gente a partir de evaluaciones erróneas.

La responsabilidad del uso adecuado de los tests sólo compete a los psicólogos, en un país como el nuestro, en vías de desarrollo, esta responsabilidad se vuelve mucho mayor. Solo a los profesionales de la conducta les compete esta tarea.

5.7 Limitaciones

La realización de esta investigación tuvo como limitantes lo siguiente:

La muestra de aplicación no fue todo lo amplia que se esperaba (10 sujetos por variable) sin embargo, se trató de cubrir esto lo mejor que se pudo.

Durante el proceso de la investigación se trató de contar con toda la información que nos permitiera entender cómo y por qué fue elaborado el Cornell Index, dado que es un instrumento creado en la década de los 40's, rastrear la información representó un problema fuerte, ya que la información sobre el Cornell Index era muy limitada, sólo hasta el final se contó con el manual original del instrumento.

Investigaciones de este tipo desafortunadamente implican problemas, como son el de tipo económico, porque al ser un instrumento de origen extranjero, la obtención de materiales originales representa gastos fuertes, gastos que muchas veces no pueden ser cubiertos en su totalidad. Esto representa una de las principales limitantes que en general se presentan al hacer investigación en Psicometría, de ahí que poca gente se interese por realizar investigación en esta área.

5.8 Sugerencias

Para posteriores investigaciones se sugiere trabajar desde el inicio con el material original del instrumento, en comunicación personal con el autor del Cornell Index, el autor de esta investigación planea trabajar con la prueba original utilizando muestras poblacionales normales y con padecimientos como los que pretende medir el instrumento, lo que se llamaría una validación cruzada (Weider, 1995).

Una segunda sugerencia sería el rescatar los reactivos que se comportaron mejor y elaborar un instrumento nuevo tomando en cuenta todos los aspectos que debe cubrir un instrumento, como son Confiabilidad y Validez adecuada.

Otra sugerencia que se hace es el de elaborar instrumentos con propósitos semejantes a los que mide el Cornell Index, pero creados a partir de la idiosincrasia de los mexicanos, serían de gran ayuda en el campo profesional.

REFERENCIAS DOCUMENTALES

REFERENCIAS DOCUMENTALES

- Aiken, L.R. (1995) **Psychological testing and assessment.** U.S.A.: Allyn and Bacon.
- Aiken, L.R. (1996) **Personality Assessment: Methods and Practices.** (2nd Ed.) Kirkland, WA: Hogrefe & Huber Publishers.
- Anastasi, A. (1988) **Psychological testing.** (6th Ed.). New York.: MacMillan Pub. Comp.
- Bischof, L. S.(1973) **Interpretación de las teorías de la personalidad.** México: Ed. Trillas..
- Brown, F.G. (1980) **Principios de la medición en psicología y educación.** México: Ed. El Manual Moderno.
- Buros, O.K. (1953) **The fourth mental measurement yearbook.** Highland Park, N.J.: Gryphon Press.
- Buros, O.K. (1959) **The fifth mental measurement yearbook.** Highland Park, N.J.: Gryphon Press.
- Buros, O.K. (1974) **Tests in print II.** Highland Park, N.J.: Gryphon Press.
- Buros, O. K. (1978) **The eighth mental measurement yearbook.** Highland Park, N.J.: Gryphon Press.
- Butcher, J.N. y Panchieri, P. (1976) **A Handbook of Cross-National MMPI. Research.** Minnesota: University of Minnesota Press.
- Cerda, E. (1957) **Validación del Cornell Index FN2 en sujetos españoles,** en "Rev. de Psic. Gral. y Aplic", vol. XII, No. 44, 752-776.
- Cerda, E. (1978) **Psicometría general.** Barcelona: Ed. Herder.
- Cerda, E.(1984) **Psicología aplicada.** (9nva. Ed.)Barcelona: Ed. Herder.

- Cueli, J. (1972) **Teorías de la personalidad**. México: Ed. Trillas.
- Dahlstrom, G.W. (1993) **Tests: pequeños ejemplos, grandes consecuencias**. *American Psychologist*, Abril 1993, pag. 393-399. (Traducción)
- Diaz-Guerrero, R. (1971) **La enseñanza de la investigación en Psicología en Iberoamérica: Un Paradigma**. *Revista Latinoamericana de Psicología*, Vol. 3, No. 1, 5-36.
- Diaz-Guerrero, R. y Salas, M. (1975) **El diferencial semántico del idioma español**. México: Ed. Trillas.
- Engler, B. (1995) **Personality Theories: An introduction (4th Ed.)** Boston, MA: Houghton Mifflin.
- Fisseni, H. J. (1987) **Psicología de la personalidad**. Barcelona: Ed. Herder.
- Geiwitz, J. (1977) **Teorías no freudianas de la personalidad**. (2da. Ed.) Madrid: Ed. Marova.
- Gregory, R. J. (1993) **Psychological Testing: history, principles, and applications**. Boston: MA. Allyn & Bacon.
- Groth-Marnat, G. (1990) **Handbook of Assessment**. (2nd. Ed.) New York: Willey.
- Guerrero, L. A. (1987) **Estudio exploratorio en un grupo de adolescentes de una colonia marginada a través de la prueba de Cornell Index**. UNAM. Tesis de Licenciatura.
- Guerrero, L.A. (1990) **Prueba de Cornell Index. Material Didáctico para las prácticas de Evaluación de la Personalidad**. Facultad de Psicología, UNAM.
- Hall, C.S. y Lindzey, G. (1984) **La teoría factorial de la personalidad**. Buenos Aires: Ed. Paidós.

- Hall, C.S. (1975) **Las grandes teorías de la personalidad**. Buenos Aires: Ed. Paidós.
- Hernández, G. L. (1995) **Ansiedad en adolescentes**. Memorias del VII Congreso Mexicano de Psicología. México
- Hernández, M.S., León, B. L., Martínez, F. J. (1985) **Lecciones sobre metodología de las Ciencias Sociales**. Ciudad Universitaria, México. Universidad Nacional Autónoma de México.
- Holtzman, H. W. (1989) **Psychology and Health Around the World**. Invited lecture at the University of Madrid, July 4, 1989.
- Kline, P. (1983) **Personality: Measurement and theory**. London: Hutchinson Pub.
- Kline, P. (1986) **A handbook of test construction: Introduction to psychometric design**. London : Methuen & Co. Ltd.
- Kline, P. (1993) **Personality: The Psychometric view**. London: Routledge.
- Kline, P. (1993) **Handbook of psychological testing**. London: Routledge.
- Lanyon, R.I. y Goodstein, L.D. (1977) **Evaluación de la personalidad**. México: Ed. El Manual Moderno.
- Magnusson, D. (1973) **Teoría de los tests**. México: Ed. Trillas.
- Mann, W. (1950) **The validation of the Cornell Index for freshmen at Michigan State College**. Abstract., *Am. Psychol.*, 5:349-350.
- Martínez, A. R. (1995) **Psicometría: Teoría de los tests psicológicos y educativos**. Madrid: Ed. Síntesis S.A.
- Marx, H.M. (1963) **Sistemas y teorías psicológicas contemporáneos**. Buenos Aires: Ed. Paidós.
- Mischel, W. (1988) **Teorías de la personalidad**. (2da. Ed.) México: Ed. Mc. Graw Hill.

- Mittelman, B.; Brodman, K. (1946) **The Cornell Indices and the Cornell Word Form: 1. Construction and Standardization.** in *Ann. N. Y. Acad. Sci.* 46: 573-7.
- Morales, M.L. (1975) **Psicometría aplicada.** México: Ed. Trillas.
- Morgan, H. y Cogger, J. W. (1975) **El manual del entrevistador.** México: El Manual Moderno.
- Murphy, K.R. & Shofer, D. (1994) **Psychological testing: Principles and applications.** (3th) New Jersey: Prentice-Hall.
- Nunnally, J.C. (1970) **Introducción a la medición psicológica.** Buenos Aires: Ed. Paidós.
- Nunnally, J.C. (1995) **Teoría psicométrica.** México: Ed. McGrawhill.
- Pervin, L. A. (1993) **Personality: theory and research.** (6th Ed.) U.S.A.: John Willey & Sons, Inc.
- Pichot, P. (1976) **Los tests mentales.** Buenos Aires: Ed. Paidós.
- Reidl, L. (1990) **Medición en ciencias sociales.** Documento inédito.
- Sattler, J. (1984) **Evaluación de la inteligencia infantil y habilidades especiales.** México: El Manual Moderno.
- Tyler, L. (1972) **Pruebas y medición en psicología.** Colombia: Prentice Hall International.
- Tyler, L. (1972) **Psicología de las diferencias humanas.** Madrid: Marova.
- Villegas, H. E. y Varela, D. R. (1992) **Análisis de confiabilidad del "Índice Cornell FN2".** Memorias del VI Congreso Mexicano y II Iberoamericano de Psicología Clínica. México.
- Weider, A.; Wechsler, D. (1946) **The Cornell Indices and the Cornell Word Form: 2. Results.** *Ann. N.Y. Acad. Sci.* 579-587

Weider, A. (1975) **The Cornell Index: Manual**. New York: Private Edition.

Weider, A & Varela, R. (1995) **Comunicación personal**

Wolff, H. (1946) **The Cornell Indices and the Cornell Word Form: 3, Application**. *Ann. N.Y. Acad. Sci.* 603-605

APENDICE

C O R N E L L I N D E X

NOMBRE _____ EDAD _____
OCUPACION _____ SEXO _____
ESCOLARIDAD _____ EDO. CIVIL _____
FECHA _____ Dx. _____

INSTRUCCIONES

TRACE UN CIRCULO ALREDEDDR DE LA PALABRA SI , CUANDO SU RESPUESTA SEA AFIRMATIVA. CUANDO SEA NEGATIVA, ENCIERRE EN EL CIRCULO LA PALABRA NO. CONTESTE TODAS LAS - PREGUNTAS QUE A CONTINUACION SE LE PRESENTAN. MARQUE LA RESPUESTA QUE LE PAREZCA MAS JUSTA.

- 1.- SI NO ¿Se ha sentido anteriormente deprimido o triste?
- 2.- SI NO ¿Frecuentemente se siente decaído?
- 3.- SI NO ¿Tiene usted mucho frío o mucho calor por ratitos?
- 4.- SI NO ¿Se ha desmayado en más de dos ocasiones en su vida?
- 5.- SI NO ¿Le dan miedo los lugares o las gentes extrañas?
- 6.- SI NO ¿Padece de mareos frecuentemente?
- 7.- SI NO ¿Se siente nervioso o temeroso cuando se le acerca un superior?
- 8.- SI NO ¿Cuando ve sangre , siente desmayarse?
- 9.- SI NO ¿Se pone nervioso cuando lo observan hacer su trabajo?
- 10.- SI NO ¿Siente miedo de estar solo?
- 11.- SI NO ¿Se siente nervioso o marcado en este momento?
- 12.- SI NO ¿Se equivoca con frecuencia cuando le dan órdenes?
- 13.- SI NO ¿Se atonta cuando tiene que actuar rápidamente?
- 14.- SI NO ¿Suda y tiembla usted durante un exámen?
- 15.- SI NO ¿Desea usted que alguien le ayude?

- 16.- SI NO ¿Tiene que hacer las cosas lentamente para estar seguro que las hace bien?
- 17.- SI NO ¿Le molesta comer en cualquier parte que no sea su casa?
- 18.- SI NO ¿Tiene usted una incontrolable necesidad de repetir la misma acción ó acto aunque le moleste?
- 19.- SI NO ¿Habitualmente le es difícil tomar una decisión por sí mismo?
- 20.- SI NO ¿Se siente a menudo triste o infeliz?
- 21.- SI NO ¿Frecuentemente desea morir?
- 22.- SI NO ¿Le daría mucho miedo quedarse sin trabajo?
- 23.- SI NO ¿El futuro le parece a usted habitualmente incierto o sin esperanza?
- 24.- SI NO ¿Es usted indiferente ante situaciones violentas?
- 25.- SI NO ¿En las reuniones le es difícil entablar conversación y siente deseos de separarse del grupo?
- 26.- SI NO ¿Desearía desentenderse de los problemas que le aquejan?
- 27.- SI NO ¿Se considera usted nervioso?
- 28.- SI NO ¿Tiene miedo sin motivo?
- 29.- SI NO ¿Le cuesta trabajo dormirse o permanecer dormido?
- 30.- SI NO ¿Se preocupa demasiado por contrariedades insignificantes?
- 31.- SI NO ¿Se siente con frecuencia preocupado?
- 32.- SI NO ¿Ha tenido alguna vez un padecimiento nervioso?
- 33.- SI NO ¿Ha estado usted anteriormente hospitalizado en alguna institución para enfermos mentales o nerviosos?
- 34.- SI NO ¿Se sofoca usted con pequeños esfuerzos?
- 35.- SI NO ¿Tiene a menudo molestias en el corazón o en el pecho?
- 36.- SI NO ¿A veces siente que se le quiere salir el corazón?
- 37.- SI NO ¿Frecuentemente le es difícil respirar o le duele la cabeza?
- 38.- SI NO ¿Padece usted de palpitaciones frecuentemente, sin ninguna razón aparente?
- 39.- SI NO ¿Algunas veces se asusta de lo que piensa?
- 40.- SI NO ¿A menudo se pone usted tembloroso?

- 41.- SI NO ¿Tiene pesadillas frecuentemente?
- 42.- SI NO ¿Se asusta con las sombras y ruidos repentinos?
- 43.- SI NO ¿Los ruidos repentinos le hacen brincar?
- 44.- SI NO ¿Se pone nervioso cuando alguien grita cerca de usted?
- 45.- SI NO ¿A menudo se siente inquieto o intranquilo?
- 46.- SI NO ¿Tiene pensamientos o ideas que le impresionan mucho y que vuelven una y otra vez a su mente?
- 47.- SI NO ¿Padece frecuentemente de dolores de cabeza?
- 48.- SI NO ¿Le sudan mucho sus manos y su cuerpo aún en época de frío?
- 49.- SI NO ¿Siente usted comezón cuando lo hacen enojar?
- 50.- SI NO ¿Tartamudea frecuentemente?
- 51.- SI NO ¿Le aparecen a veces tics nerviosos en la cabeza, ojos o en cualquier parte del cuerpo?
- 52.- SI NO ¿Se orinaba usted en la cama entre los 8 y los 14 años de edad?
- 53.- SI NO ¿Padece de manos y pies fríos?
- 54.- SI NO ¿Sufre usted de asma?
- 55.- SI NO ¿En la noche moja usted la cama?
- 56.- SI NO ¿Es usted sonámbulo?
- 57.- SI NO ¿Ha tenido ataques o convulsiones?
- 58.- SI NO ¿Tiene dolores de espalda que le impiden seguir trabajando?
- 59.- SI NO ¿Algunas veces le duelen tanto los ojos que no puede ver?
- 60.- SI NO ¿A menudo se siente en muy mal estado físico?
- 61.- SI NO ¿Tiene usted achaques o dolores tan fuertes que le impiden realizar su quehacer?
- 62.- SI NO ¿Tiene temporadas de agotamiento o de fatiga?
- 63.- SI NO Se acaba usted mismo con preocupaciones acerca de su salud?
- 64.- SI NO ¿Siente usted debilidad o dolores de pies todos los días?
- 65.- SI NO ¿Al levantarse por la mañana se siente usted cansado?
- 66.- SI NO ¿Cuando tiene dolor de cabeza, o le laten las sienas, se le dificulta hacer su trabajo?
- 67.- SI NO ¿Frecuentemente se enferma y se siente decaído?

- 68.- SI NO ¿Se siente tan agotado o fatigado que no puede ni comer?
- 69.- SI NO ¿Tiene mal apetito?
- 70.- SI NO ¿Sufre constantemente de estreñimiento o diarrea?
- 71.- SI NO ¿Muy a menudo siente usted el estómago revuelto?
72. SI NO ¿Padece frecuentemente de ascos y vómitos?
- 73.- SI NO ¿Sufre de indigestión a menudo?
- 74.- SI NO ¿Frecuentemente tiene dolores de estómago o enfermedades del mismo?
- 75.- SI NO ¿Trabajan mal su estómago y sus intestinos?
- 76.- SI NO ¿Cuando acaba de comer, siente tan fuertes dolores de estómago que lo hacen doblarse?
- 77.- SI NO ¿Usualmente tiene molestias al digerir la comida?
- 78.- SI NO ¿Frecuentemente sufre de malestar porque se llena de gases?
- 79.- SI NO ¿Le ha dicho algún médico que tiene úlcera en el estómago?
- 80.- SI NO ¿Generalmente la gente lo mal interpreta?
- 81.- SI NO ¿Ha tenido la sensación de ser observado mientras trabajaba?
- 82.- SI NO ¿Generalmente lo tratan mal las otras personas?
- 83.- SI NO ¿Ha tenido la sensación de que la gente lo observa o habla de usted en la calle?
- 84.- SI NO ¿Generalmente los demás abusan de su buena fé?
- 85.- SI NO ¿Es usted tímido o sensible?
- 86.- SI NO ¿Se irrita o se inquieta fácilmente?
- 87.- SI NO ¿Se le dificulta hacer amistades?
- 88.- SI NO ¿Estallarfa si no se controla constantemente?
- 89.- SI NO ¿Alguna vez lo enviaron a una escuela correccional?
- 90.- SI NO ¿Ha tenido alguna vez serios problemas o perdido su empleo por beber?
- 91.- SI NO ¿Ha sido arrestado más de tres veces?
- 92.- SI NO ¿Ha tomado alguna vez drogas, y se le ha hecho costumbre?
- 93.- SI NO ¿Siente usted que sus enemigos se ensañan en perseguirlo?
- 94.- SI NO ¿Le molesta que alguien le diga lo qué tiene que hacer?

- 95.- SI NO ¿Frecuentemente ahoga sus penas bebiendo?
- 96.- SI NO ¿Hace cosas de repente y sin pensarlas?
- 97.- SI NO ¿Se siente habitualmente engañado por la gente o por alguien en particular?
- 98.- SI NO ¿Se enoja si no puede obtener las cosas que quiere de inmediato?
- 99.- SI NO ¿Le desagrada el sexo opuesto?
- 100.- SI NO ¿Tiene que estar siempre alerta con sus amigos o defenderse de ellos?
- 101.- SI NO ¿Frecuentemente tiene arranques de enojo?