

102
24.



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO.**

FACULTAD DE PSICOLOGIA

ESCALAS DE VALIDEZ DEL 16 PF DE CATELL Y LA
ESCALA DE DISTORSION MOTIVACIONAL (DM)
DE WINDER, O'DELL Y KARSON

T E S I S A
PARA OPTAR POR EL GRADO DE
LICENCIADO EN PSICOLOGIA
QUE PRESENTA
GUADALUPE MARCELA FUENTES GARCIA

DIRECTORA
LIC. MA. ENEDINA VILLEGAS HERNANDEZ

CIUDAD UNIVERSITARIA, MEXICO, D. F.

1997

TESIS CON
FALLA DE ORIGEN



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

DEDICATORIAS

A mi hijo, Alejandro D'Luna, quien ha iluminado mi vida desde antes de su nacimiento.

A mis padres, Dn. Carlos Fuentes y Dña. Trinidad García de Fuentes, Q.E.P.D, a quienes debia este trabajo.

A todos mis sobrinas y sobrinos, en especial a Mara Carolina. Q.E.P.D.

A Dn. Crisóforo Carrillo, quien me hizo concebir un sueño grandioso: el de su amor.

AGRADECIMIENTOS

A toda la comunidad universitaria de la UNAM. A todas y cada una de las personas que de alguna manera contribuyeron para mi formación y para la realización de este trabajo.

A la División de Educación Continua de la Facultad de Psicología de la UNAM, en especial al Programa de Apoyo a la Titulación.

A la maestra Enedina Villegas, por la dirección de mi tesina pero sobre todo, por su trabajo como maestra y formadora de generaciones de psicólogos diferenciales.

A mis sinodales y a todos los maestros de la Facultad de Psicología de la UNAM.

A las amigas y amigos que me alentaron y ayudaron y en especial, a mi hijo, por su valiosísimo, cotidiano e incondicional apoyo.

INDICE

INTRODUCCION

CAPITULO 1

LAS TEORIAS FACTORIALES, EYSENCK, GUILFORD Y CATELL

1.1 LAS TEORIAS DE LA PERSONALIDAD	P 1
1.1.2 LA PSICOLOGIA DE LOS RASGOS	P 4
1.1.3 EL ANALISIS FACTORIAL	P 10
1.2 LA TEORIA DE EYSENCK	P 14
1.3 LA TEORIA DE GUILFORD	P 18
1.4 LA TEORIA DE CATELL	P 21
1.5 INVENTARIOS DE PERSONALIDAD IMPORTANTES	
1.5.1 EL CUESTIONARIO 16 PF DE CATELL	P 32
1.5.2 EL INVENTARIO MULTIFASICO DE LA PERSONALIDAD MINNESOTA MMPI	P 39

CAPITULO 2

PRINCIPALES ENFOQUES PARA LA EVALUACION DE LA PERSONALIDAD

2.1 LA EVALUACION PSICOLOGICA	P 43
2.1.1 MODELOS DE EVALUACION	P 44
2.2 EL ENFOQUE PROYECTIVO	P 48
2.3 EL ENFOQUE PSICOMETRICO	P 62
2.3.1 LOS INVENTARIOS DE PERSONALIDAD	P 77

CAPITULO 3

EL FALSEAMIENTO DE LAS RESPUESTAS EN INVENTARIOS DE PERSONALIDAD

3.1 EL ERROR DE MEDIDA	P 85
3.2 EL FALSEAMIENTO DE LAS RESPUESTAS	P 87
3.2.1 TIPOS DE FALSEAMIENTO	P 93
3.3 LAS ESCALAS DE VALIDEZ	P 98
3.3.1 LOS TESTS DE INTEGRIDAD	P 99
3.3.2 ESCALAS DE VALIDEZ DEL MMPI	P 101
3.3.3 ESCALAS DE VALIDEZ DEL 16 PF-A	P 106
3.3.4 ESCALAS DE VALIDEZ DEL 16 PF-5	P 113

CAPITULO 4

LA ESCALA DE DISTORSION MOTIVACIONAL DM DE WINDER, O'DELL Y KARSON

4.1 ANTECEDENTES	P 115
4.1.1 DESCRIPCION DEL PROCEDIMIENTO	P 116
4.2 EL ENFOQUE CORRECTOR DE KRUG	P 120
4.2.1 UNA APLICACION DEL METODO DE KRUG	P 122
4.3 UNA ESCALA DE DISTORSION MOTIVACIONAL EN ESPAÑA	P 123

CONCLUSIONES

1. LA EVALUACION DE PERSONAL EL MEXICO	P 127
2. LA EVALUACION COMPUTARIZADA	P 131
3. CONCLUSIONES	P 135
4. LIMITACIONES Y SUGERENCIAS DE INVESTIGACION	P 136

BIBLIOGRAFIA

P 138

INTRODUCCION

La preocupación principal de la Psicología debe ser el comprender a la Persona. Independientemente del enfoque teórico o metodológico, de los modelos o de las corrientes que el psicólogo prefiera, es la Persona el inicio y debe ser el fin de esta ciencia.

La Psicología pretende un conocimiento de las personas que permita predecir su conducta futura, pero lo pretende desde una perspectiva científica (Fernandez-Ballesteros, 1992) Para estudiar a la persona, o el constructo más importante que es la Personalidad, han surgido innumerables teorías. Cuando la Psicología se convierte en una ciencia aplicada por las necesidades de supervivencia del individuo y de la sociedad es cuando el psicólogo requiere elegir tal o cual corriente, buscando el enfoque teórico y la metodología que mejor resuelva el problema particular que esta enfrentando.

Todos los días y a todas horas, unas personas toman decisiones sobre otras y en este proceso interviene una gran cantidad de prejuicios y sesgos personales. (Colom, 1995) El proceso de evaluación de las personas está presente en cualquier interacción social y es especialmente importante en los grupos que pretenden un fin productivo. Las organizaciones empresariales tienen necesidad de contar con los individuos que mejor colaboren con la consecución de sus fines. Para evaluar personal se cuenta con diferentes procedimientos y técnicas, desde un método intuitivo en donde un microempresario se basa en la observación o en algún tipo de interacción verbal para decidir la contratación de su único empleado hasta la gran organización que cuenta con las técnicas más avanzadas y el personal más especializado.

Muchos son los métodos existentes para evaluar a las personas. Los tests psicológicos, bien contruidos y convenientemente manejados en todo el proceso de evaluación, pueden reducir los sesgos y prejuicios, dado que las personas somos malos jueces de otras personas. (Colom, 1995) La evaluación de la personalidad es el esfuerzo que la Psicología ha realizado para describir y usar predictivamente las consistencias del comportamiento de una persona particular, de una manera objetiva y científica. (Fernandez-Ballesteros, 1992)

Con la convicción de que las personas son diferentes entre si en un gran número de características y que esas diferencias pueden ser medidas, comienza el estudio de las diferencias individuales. La Psicología Diferencial intenta proporcionar una serie de respuestas usando instrumentos de medida. La Psicología Diferencial no pierde de vista al individuo como un ser integral aunque como método de estudio parta de los rasgos como elementos básicos.

Investigadores como Eysenck, Guilford y Catell, entre otros, han dedicado gran parte de su vida a la investigación de los rasgos, en un intento por condensar matemáticamente dichas cualidades a partir del comportamiento concreto de personas concretas ante instrumentos concretos. (Colom, 1995) Debido a que estos investigadores han utilizado modelos factoriales para el estudio de la personalidad, han sido conocidos como factorialistas. Los inventarios factoriales son los instrumentos derivados de sus investigaciones, entre los que se encuentra el Cuestionario 16 PF de Catell. En el primer capítulo se realiza una somera revisión de las teorías de la personalidad, se expone la Psicología de los rasgos, se explica el método factorial, se revisan las teorías de los principales factorialistas y se termina con una descripción del Cuestionario 16 PF de Catell, que es el instrumento objeto de este trabajo.

En la evaluación de la personalidad existen dos enfoques claramente diferenciados: el enfoque psicométrico y el enfoque proyectivo. Ambos se basan en distintos modelos de evaluación: manejan diferentes variables, utilizan diferentes métodos y técnicas y se orientan a distintos objetivos. Asimismo, han dado lugar a diferentes instrumentos o técnicas de evaluación. Los inventarios son los instrumentos más representativos del enfoque psicométrico. Estos temas son expuestos en el segundo capítulo.

En mi práctica profesional, llevada a cabo primordialmente en el campo de la selección de personal, ha sido una preocupación constante el realizar una evaluación precisa, objetiva, confiable y que permita efectivamente el predecir algunos rasgos o tendencias de personalidad importantes en el desempeño del puesto. He considerado que la evaluación es una gran responsabilidad con las empresas y con las personas. Me ha interesado especialmente la tendencia de los evaluados a falsear sus las respuestas a los inventarios de personalidad cuando saben que una decisión será tomada con base en los resultados de tal evaluación.

El falseamiento de las respuestas ante instrumentos de evaluación ha sido objeto de estudio de los investigadores, desde aquellos pioneros que se percataron que la gente fingía enfermedades mentales con algún propósito particular. Los inventarios de personalidad, que se basan en una autodescripción de parte del evaluado, son especialmente susceptibles al falseamiento; Furham (1992) considera que ésta es una de las principales desventajas de los tests de personalidad en el contexto laboral. Para contrarrestar esta desventaja de los inventarios, se han construido escalas adicionales, denominadas escalas de validez, para identificar el eventual falseamiento, el tipo y la magnitud del mismo y de esta forma surgen escalas tan famosas y útiles como las escalas de validez del Inventario Multifásico de la Personalidad Minnesota MMPI.

En el capítulo 3 se exponen los temas referentes al falseamiento y las escalas de validez del MMPI y del Cuestionario 16 PF de Catell; la escala que identifica la tendencia a aparecer mejor o de Distorsión Motivacional, la de distorsión negativa y la de azar.

En el capítulo 4 se explica el procedimiento seguido por Winder, Karson y O'Dell (1975) para derivar la escala de Distorsión Motivacional, la escala de distorsión negativa y la escala de respuestas al azar del Cuestionario 16 PF de Catell. Asimismo, se refieren algunas investigaciones importantes sobre el tema.

En la parte final, se exponen algunas consideraciones sobre la evaluación psicológica de personal en México y otras acerca de la evaluación psicológica computarizada para finalizar con la presentación de las conclusiones, las limitaciones de este trabajo así como las sugerencias para futuras investigaciones.

Este trabajo es un intento, muy limitado, de contestar a las múltiples preguntas que se hacen tanto los psicólogos en la evaluación de personal como los evaluados y los referentes o clientes organizacionales, en relación a la utilidad de los tests, a su confiabilidad, a su susceptibilidad al falseamiento, a su poder de predicción, etc.

Algunas consideraciones lingüísticas

En este trabajo se han utilizado algunos términos en inglés como *test*, *assessment center* y otros. Gramaticalmente es un barbarismo anglicano el usar palabras inglesas en un texto en español pero se prefiere este error a una imprecisión de significado (Uribe Prado, 1996). Se optó por esta modalidad por la razón del uso extendido de tales términos en la literatura técnica en todos los idiomas. También se ha preferido usar el nombre del Cuestionario de Catell con las siglas en inglés, 16 PF (personality factors) en lugar de su traducción al español, 16 FP (factores de personalidad), por la amplitud de uso del primero. Tales términos se han usado en este trabajo, sin cursivas.

A propósito, el término *TEST* es un vocablo inglés que significa prueba y proviene del vocablo latino *testis-testis* que significa testigo, testimoniar y que también se traduce como declaración o prueba (Diccionario ilustrado Vox, latino-español, 1990). El Diccionario Enciclopédico Ilustrado (1979), en español, incluye la palabra *test* y la define como prueba, ensayo, análisis, examen.

Finalmente, los factorialistas acostumbran asignar un nombre a los rasgos o dimensiones de personalidad identificados por análisis factorial. Catell con frecuencia uso neologismos para sus factores que se usan de la misma forma en cualquier idioma. En este trabajo se han respetado los términos de Catell o la traducción que hace cada investigador de tales términos.

CAPITULO 1.

LAS TEORIAS FACTORIALISTAS: EYSENCK, GUILFORD Y CATELL.

Para entrar en el tema de este capítulo, se presentará información acerca de las teorías de la personalidad, de la Psicología de los rasgos y del análisis factorial.

1.1 LAS TEORIAS DE LA PERSONALIDAD

Una teoría es una explicación, un "conocimiento especulativo considerado con independencia de sus posibles aplicaciones, una suposición o hipótesis, una ley o sistema de leyes que se deducen de la observación de ciertos fenómenos y sirven para relacionarlos o explicarlos científicamente" (Diccionario enciclopédico ilustrado, 1979)

Hall y Lindzey (1991), los autores que han realizado una gran obra en el estudio de las teorías de la personalidad, refieren que una teoría es un conjunto de convenciones creadas por el teórico y que los datos surgidos de una investigación pueden ser incorporados a uno entre muchos esquemas teóricos diferentes. Consideran impropio asignarle a una teoría atributos de verdad o falsedad y prefieren calificar su utilidad o inutilidad para generar predicciones o proposiciones verificables acerca de acontecimientos pertinentes.

Pervin (1970, 1990) concuerda con ellos en cuanto a la utilidad de las teorías y agrega que puede aceptarse aquella que cumpla con criterios de globalidad, parsimonia y relevancia, es decir, que "una teoría de personalidad vale si con pocos elementos explica fenómenos variados y permite generar nuevos conocimientos" (en Fernández-Ballesteros, 1992).

La Psicología pretende un conocimiento de las personas que permita predecir su conducta futura pero lo pretende desde una perspectiva científica. El concepto o constructo de *personalidad* es el que mejor cubre esa necesidad de conocimiento del sujeto. En la historia de la Psicología ha habido innumerables intentos de definir la personalidad e incluso, algunas corrientes han negado su existencia quizá por las dificultades científicas de abordar su investigación. La mayoría de las definiciones concuerdan en considerar a la personalidad como "aquello que permite prever las tendencias estables de una persona a comportarse de determinada manera en diferentes situaciones". La evaluación de la personalidad reúne los esfuerzos para describir y usar predictivamente las consistencias conductuales de una persona determinada, de una manera objetiva y científica (Fernández-Ballesteros 1992)

El término *personalidad* ha sido objeto de un gran número de definiciones, Allport en 1937 ya había encontrado casi cincuenta. Hall y Lindzey (1991) consideran que ninguna definición esencial de la personalidad puede aplicarse en forma general pero logran identificar siete tipos de definiciones de personalidad, parafraseando a Allport:

-la definición biosocial que equipara la personalidad con el valor de estímulo social del individuo;

-la definición biofísica que indica un aspecto orgánico de la personalidad y que la explica en función de las características o cualidades del sujeto;

-la definición miscelánea la cual se vale de la enumeración de los conceptos fundamentales para la descripción de la persona;

-la definición integradora u organizadora que considera a la personalidad como la organización o pauta que da coherencia a las diversas conductas del sujeto;

-la definición que enfatiza el ajuste o que considera a la personalidad como intermedia en los esfuerzos que realiza el sujeto para lograr precisamente su adaptación al medio;

-la que considera los aspectos individuales o singulares de la conducta, los rasgos distintivos de la persona que la diferencian de las demás.

-por último, se mencionan a los teóricos que consideran a la personalidad como la esencia del hombre, lo más representativo de cada persona o, en términos de Allport, "la personalidad consiste en aquello que un hombre realmente es"

Hall y Lindzey (1991) reconocen cuatro importantes corrientes que han ejercido influencia sobre las teorías de la personalidad en épocas recientes.

a) la derivada de la observación clínica que inicia con Charcot y Janet y que tiene como importantes representantes, entre otros, a Freud, Jung y McDougall.

b) la que surge con los técnicos gualtistas y William Stern, quienes le otorgaban máxima importancia a la unidad de la conducta en contra de la concepción del estudio de la misma por segmentos o pequeños elementos. Este concepto se maneja en algunas de las actuales teorías de la personalidad.

c) la representada por el enfoque de la Psicología experimental en general y la teoría del aprendizaje en particular que da importancia relevante a la investigación empírica controlada y ha permitido comprender la naturaleza de la elaboración de una teoría.

d) El último enfoque es el derivado de la tradición psicométrica que se centra primordialmente en el estudio y en la medición de las diferencias individuales. Estos investigadores han desarrollado la metodología de la medición y el análisis cuantitativo de los datos.

Colom (1995) refiere una importante revisión de las teorías de la personalidad llevada a cabo por Pederson, Allen, Laue y Johnson (1992), del Laboratorio Armstrong de la Fuerza Aérea de EE.UU. El propósito del estudio era encontrar aquella teoría que tuviera mayor valor como respaldo de la evaluación de las diferencias individuales. En la revisión incluyeron teorías estructurales de investigadores diferencialistas y otras teorías no estructurales.

Las fases seguidas por Pederson y sus colegas en esta revisión fueron: a) Identificación de las teorías. b) Valoración de las teorías. c) Selección de la teoría que mejor se ajustara a los criterios previstos. Las teorías revisadas fueron las siguientes:

ESTRUCTURALES

Teoría de Eysenck

Teoría de Cinco factores de McCrae y Costa

Teoría de 16 factores de Cattell

Teoría *Circumplex* de Wiggins y Broughton (1985)

Teoría socioanalítica de Hogan (1983)

Teoría de los tipos psicológicos de Jung (1923)

Teoría temperamental de Buss y Plomin

Teoría de la Expresividad-instrumentalidad de Spence (1983)

Teoría de las necesidades de Murray (1938)

NO ESTRUCTURALES.

Teoría de la autoeficacia de Bandura (1986)

Teoría situacional de Mischel (1968)

Teoría del esquema personal de Markus (1977)

Teoría sociocognitiva de Cantor y Kihlstrom (1987)

Los criterios de valoración de las teorías fueron los siguientes.

Criterios generales.

1. Utilidad para la predicción. Una teoría debe proporcionar un mecanismo sistemático para facilitar el descubrimiento de nuevas relaciones empíricas entre variables, debe sugerir predicciones concretas que permitan estudios empíricos específicos.
2. Supuestos explícitos. Estos deben indicar de modo explícito los eventos empíricos con los que se relaciona la teoría.
3. Predicciones contrastables. Las definiciones que recoja la teoría deben ser explícitas y prácticas de modo que los conceptos teóricos se puedan traducir en predicciones medibles que permitan generar datos verificables.
4. Incorporación de los resultados conocidos. Una teoría debe organizar e integrar todo lo que se sabe sobre el área de conocimiento en cuestión.
5. Predicciones confirmables. La evaluación de la validez empírica resulta prematura hasta no desarrollar de modo completo la teoría.
6. Parsimonia. La teoría más sencilla será la mejor.

Criterios operativos.

1. La teoría se debe centrar en adultos normales
2. Debe ser compatible con medidas de aptitud
3. Debe haber ajuste ambiente-persona
4. Debe incluir la motivación de logro.
5. Debe ser sensible para detectar diferencias individuales independientemente de la homogeneidad de la muestra.
6. Sesgo de respuestas. Debe reducir o controlar la tendencia a la condescendencia y el falseamiento, fundamentalmente.
7. Riesgo legal. La teoría debe permitir ajustarse a distintas poblaciones de personas.

RESULTADOS. De acuerdo con los criterios mencionados, estos investigadores descartaron las siguientes: la teoría de Catell, la de Buss y Plomin, la de expresividad-instrumentalidad y la de Hogan porque no se han replicado satisfactoriamente. La de Eysenck y la de Wiggins no son teorías completas. La teoría de Murray y la de Jung se descartaron porque no incluyen una clasificación empírica. Las teorías de Bandura y de Mischel no son realmente teorías de la personalidad sino del comportamiento, de modo que no se han operativizado convenientemente; lo mismo sucede con las teorías de Markus y la de Cantor y Kihlstrom.

En resumen, la única teoría que resultó ilesa fue la de los cinco factores. Esta taxonomía pentafactorial parece la idónea según los criterios previamente fijados. Se puede confiar en ella para desarrollar medidas completas de las diferencias individuales en personalidad. En su informe, estos investigadores mencionan que el marco de referencia del modelo pentafactorial de la personalidad procede de un proyecto de investigación financiado también por la Fuerza Aérea de EEUU. en el que participaron Tupes, 1957 y Tupes y Christal, 1961.

Este estudio reviste especial importancia por ser una revisión de las teorías de personalidad encaminado a un propósito de evaluación en un escenario laboral de importancia estratégica en el ámbito mundial.

1.1.2 LA PSICOLOGIA DE LOS RASGOS

La Psicología de los rasgos ha logrado un conjunto de suposiciones básicas sobre la conducta humana que le confieren la categoría de teoría más que de una técnica para poner a prueba los conceptos de otras teorías como la consideran sus detractores. Debido principalmente a la metodología factorial, la psicología de los rasgos ha llegado a adquirir una posición relativamente autónoma entre otras teorías. (Buss y Poley, 1979)

Los psicólogos de los rasgos parten del principio de que la gente difiere entre si en varias características o rasgos, que estas diferencias pueden ser medidas y que cada individuo puede ser ubicado en un continuo. Se manejan escalas unipolares para rasgos de capacidad y escalas bipolares para rasgos de temperamento, motivación, etc.

El término rasgo se refiere a las diferencias constantes entre el comportamiento o las características de dos o más personas. Guilford considera que un rasgo es "cualquier forma discernible y relativamente constante en que un individuo se diferencia de otro" pero también le confiere un valor explicativo de la conducta y lo considera una razón hipotética que justifica la existencia de diferencias individuales persistentes. (Mischel, 1988)

Al estudiar y evaluar estos rasgos mediante un procedimiento estadístico llamado análisis factorial, los psicólogos de los rasgos se convierten en factorialistas. Los factorialistas buscan "descomponer" la personalidad en elementos denominados rasgos o factores lo cual facilita su estudio y permite una descripción particular de cada individuo

El método factorialista identifica matemáticamente los rasgos y los convierte en factores: los factores son construcciones hipotéticas o variables intervectoras que se pueden relacionar tanto al estímulo (el reactivo de una prueba, una situación) como a las respuestas. Si es posible reconocer e identificar un factor en una amplia variedad de situaciones y condiciones se piensa que representa una unidad funcional. (Cueli y Reidl, 1990) Los factores son identificados mediante el método de análisis factorial del que se expondrá una breve explicación posteriormente.

Los conceptos de Gordon Allport (1937 en Mischel, 1988) sobre los rasgos tuvieron una importante influencia sobre los factorialistas. Allport considera que los rasgos son tendencias o predisposiciones determinantes a emitir una respuesta y las realidades últimas de la organización psicológica. Define el rasgo como "un sistema neuropsíquico, peculiar del individuo, generalizado y focalizado, que tiene la capacidad de producir muchos estímulos funcionalmente independientes así como el iniciar y dirigir formas equivalentes (consistentes en forma significativa) de comportamiento adaptativo y expresivo."

Allport clasificó los rasgos en cardinales, centrales y secundarios en función de la generalización de estos rasgos en cada persona. Consideraba que no existen dos individuos iguales y que la conducta está determinada por una estructura de rasgo particular. Sus teorías contribuyeron a la teoría del rasgo pero por otro lado, criticó los métodos estadísticos y las estrategias cuantitativas de investigación de otros psicólogos de los rasgos; los factorialistas (Mischel, 1988)

Por su parte, Buss (1989 en Fernández-Ballesteros, 1992) considera que "si hay algo que pueda ser una especialidad llamada personalidad, su única y definitiva característica es el rasgo"

A diferencia de aquellas teorías de la personalidad que subrayan la importancia de las experiencias de la primera infancia, los psicólogos de los rasgos se concentran en la forma en que las personalidades adultas son distintas entre sí. Primariamente se enfocan al estudio de la conducta normal aunque la metodología es aplicable también a conducta anormal y parte de las investigaciones se refieren a este campo. También han realizado investigación sobre aspectos del desarrollo y han generado metodología para abordar este aspecto.

Los investigadores de la psicología de los rasgos no se han puesto de acuerdo en el empleo del término *personalidad* y los campos de estudio. Cattell (1982) utiliza la palabra *personalidad* para referirse a la totalidad de la conducta e incluye una amplia gama de rasgos como las capacidades, el temperamento, los motivos, los estados de ánimo, los intereses, etc.

Guilford aplica el término *personalidad* en un sentido aún más amplio que incluye rasgos morfológicos, fisiológicos, además de rasgos de temperamento, aptitudes, etc. Eysenck (1967) parece ubicar el término *personalidad* como sinónimo de temperamento pero bajo este rubro realizó estudios asociando las diferencias individuales de temperamento con las diferencias en medidas fisiológicas, el condicionamiento, la percepción, etc. (Buss y Poley, 1979)

Buss y Poley (1979) han resumido algunas de las suposiciones y características sobresalientes del enfoque de la Psicología de los Rasgos.

1. La psicología de los rasgos cuenta con una perspectiva ecléctica en relación a la influencia que determina los rasgos, desde el componente genético hasta el condicionamiento clásico lo cual permite incluir en su campo de estudio a los descubrimientos de otras teorías. Proporciona una descripción amplia de dimensiones de temperamento, capacidades, motivaciones, etc.

2. El enfoque de los rasgos y su metodología permite utilizar estos ya sea como variable dependiente o como variable independiente. En evaluación de personal se utilizan los rasgos como variables dependientes. Se manejan como variables independientes en los casos en que el interés radica en el papel de los constructos disposicionales en relación con las conductas específicas importantes desde el punto de vista social, por ejemplo.

3. Los psicólogos de los rasgos consideran a estos como una unidad fundamental en el desarrollo de los individuos y la maduración por sí misma representa una fuerza en el desenvolvimiento de la complejidad organizada de la conducta. Los efectos del ambiente para moderar el crecimiento y el desarrollo de los rasgos no se niegan pero los investigadores de los rasgos no consideran específicamente etapas en el desarrollo de los rasgos. Tampoco hay evidencia que contradiga necesariamente la teoría de las etapas.

4. La psicología de los rasgos considera que en los primeros años de vida del individuo surge una estructura de la personalidad sumamente compleja pero no se hace un especial hincapié en la importancia del ambiente para moldear esta estructura. Solamente llegan a la conclusión de que los contactos ambientales tempranos tienen una mayor importancia para la producción de tal complejidad sin llegar a considerar a tal influencia como crucial, como en el caso de otras teorías.

5. En la teoría de los rasgos se le da una importancia específica a los determinantes genéticos como una faceta de las influencias biológicas. Se considera que las influencias genéticas se encuentran en la base de las diferencias individuales o al menos en parte de ellas.

6. La psicología de los rasgos conceptualiza a éstos como unidades de análisis pero no excluye necesariamente a los conceptos más estrechos o más amplios para el análisis de la conducta.

7. La psicología de los rasgos le otorga una importancia medular a la psicometría o la medición psicológica más que cualquier otra teoría de la personalidad. Plantea que todos los aspectos cuantificables de la personalidad se pueden estudiar desde esta perspectiva. Esta amplitud de estudio proporciona la posibilidad para interrelacionar distintas facetas de diferencias individuales como capacidades-temperamento en un intento por integrar lo que se ha considerado como aspectos independientes del sistema total de la personalidad.

8. La psicología de los rasgos adopta una posición relativamente neutra en cuanto a considerar temas filosóficos más amplios como la "bondad" o la libertad en la naturaleza humana. Considera a los rasgos como positivos o negativos en función de los valores sociales. Esto no incluye a los rasgos que se miden unipolarmente -los de capacidad- en que hay poco desacuerdo sobre el extremo de la dimensión a la que se le atribuye mayor valor.

9. Las teorías de la personalidad pueden tener implicaciones sociales que usualmente se traducen en varias ideas utópicas y de control social. La psicología de los rasgos considera que la sociedad estará en condiciones de trazar la dirección que deba tomar la evolución humana hasta que se entienda la complejidad de las diferencias individuales y de sus orígenes. La supervivencia y a adaptabilidad de cualquier especie depende de la variación. El hincapié en las diferencias individuales no implica desigualdad en un sentido moral o ético.

¿CUANTOS FACTORES?

Una de las principales preocupaciones de los factorialistas ha sido el definir el número de factores de la personalidad. En el campo del temperamento, Eysenck inicialmente consideró dos rasgos y luego aumentó otro por lo que su modelo solamente considera tres superfactores. El modelo de Guilford contemplaba alrededor de trece rasgos. Cattell refiere 15 rasgos primarios y ocho secundarios aunque luego agregó los "siete rasgos ausentes" por lo que habla de 22 rasgos primarios en personalidad adulta pues en niños y en adolescentes el número es menor.

"LOS CINCO GRANDES"

Otros investigadores han insistido en que cinco rasgos son los definitorios de la personalidad, al parecer a partir de una idea de McDougall (1932 en Fernandez-Ballesteros, 1992) que sugiere la existencia de cinco elementos explicativos de la personalidad: intelecto, carácter, temperamento, disposición y templanza.

Estos elementos tienen cierta similitud con los hallazgos de los factorialistas. Entre ellos están Tupes y Christal (1961), Goldberg, 1980,1981, y Costa y Mc Rae, 1986 (en Aiken, 1996). Otros factorialistas que han seguido el modelo de "los cinco grandes" son Norman, 1963, Borgatta, 1964, Brand 1984, Hogan 1986, Lorr, 1986 y Digman,1988 (en Morris, 1992) Algunos de estos autores han desarrollado inventarios pero la mayoría de ellos no han sido traducidos al español ni adaptados a México. Goldberg (1981 en Colom,1995) sugirió que *cualquier* modelo que pretenda estructurar las diferencias individuales llegará a cinco dimensiones, en un determinado nivel de abstracción.

Norman, uno de los autores más conocido en este enfoque, dió la primera denominación de "los cinco grandes" (the Bigfive) a estos factores. La siguiente es su estructura factorial de la personalidad con un ejemplo de los adjetivos característicos pues cada factor agrupa de 30 a 40 adjetivos.

Factores de Norman	Ejemplo polo A	Ejemplo polo B
I. Extraversión	hablador	silencioso
II. Conformidad	bondadoso	intable
III. Conciencia	meticuloso	descuidado
IV. Estabilidad emocional	relajado	ansioso
V. Cultura	creativo	torpe

(Norman, 1963 en Fernandez-Ballesteros, 1992)

Cada autor ha generado su propio instrumento para evaluar estos factores. En nuestro idioma y adaptados a España, existen dos pruebas dentro de este modelo: El cuestionario *Big Five*, de Caprara, Barbaranelli y Borgogni y el Cuestionario 16 PF-5 de Catell et al. que evalúa cinco dimensiones globales o factores secundarios además de los 16 factores primarios. (Catálogo de pruebas de Tea Ediciones)

Este modelo de cinco factores para englobar los rasgos más relevantes de la personalidad ha tenido bastantes seguidores. La referida investigación de Pedersen et al (1992 en Colom, 1995) para la Fuerza Aérea de EEUU, que analizó las principales teorías de la personalidad, llegó a la conclusión que el modelo pentafactorial era el idóneo para la investigación y evaluación de las diferencias individuales y ellos eligieron específicamente a la teoría de McCrae y Costa.

CRITICAS A LOS FACTORIALISTAS

Muchos de los investigadores de la personalidad no consideran que el análisis factorial, pese a sus avances matemáticos, produzca verdaderas dimensiones de la personalidad en algún sentido sustancial. Se ha encontrado que la validez con relación a criterios de los inventarios de la personalidad construidos sobre la base del análisis factorial tiende a ser baja o desconocida y por lo mismo, los inventarios factoriales son menos útiles al hacer predicciones y tomar decisiones conductuales como diagnóstico, intervención, capacitación, etc en contextos aplicados, que los inventarios con validez de contenido y codificación de criterios. (Aiken, 1996)

Una de las críticas que se le hacen a estos investigadores alega que los factorialistas "crean sistemas de artefactos sin una verdadera relación con algún individuo singular y, por lo tanto, distorsionan y falsean la realidad" (Hall y Lindzey, 1984) pero es precisamente esa característica lo que le confiere uno de sus grandes aciertos su utilidad en la evaluación de grupos numerosos de individuos, la obtención de medidas que caractericen a un grupo en especial, la evaluación de personalidad preliminar rápida, la comparación intragrupos en un grupo y la comparación entre grupos e individuos, entre otras aplicaciones, sin que se limite su aplicación al tratar a un individuo en particular

Otra crítica sostiene que del análisis factorial se obtiene lo que se ha puesto en él: que el investigador transfiere su subjetividad o su intuición al decidir que dimensiones han de introducirse en la matriz de correlaciones y que esta misma subjetividad esta presente al interpretar los hallazgos factoriales. Esta crítica no es totalmente válida pues Eysenck, por ejemplo, utilizó el método de análisis de criterio que subraya la necesidad de disponer de hipótesis acerca de la naturaleza del factor psicológico antes de enfrentar su estudio. Catell, por su lado, partió de la esfera de la personalidad del lenguaje para sus estudios factoriales de temperamento. (Hall y Lindzey, 1984)

Otro aspecto que ha sido blanco de críticas, también mencionado por Hall y Lindzey, (1984) han sido las discrepancias entre los métodos de análisis factorial utilizados por los distintos factorialistas. Ni siquiera existe un acuerdo en cuanto a los méritos de la rotación oblicua sobre la ortogonal. Incluso factorialistas de la talla de Eysenck y Catell, ambos graduados de la Universidad de Londres, han criticado entre si los resultados de sus análisis factoriales porque uno usa la rotación ortogonal y el otro la oblicua (Furham, 1992)

Muchos psicólogos incluso no les confieren rango de teorías a las teorías factorialistas pues consideran que son sistemas que no profundizan en los procesos evolutivos, en la exposición de los supuestos sobre la conducta, etc. (Hall y Lindsey, 1984) En este sentido, los autores que se mencionan en este trabajo ubican sus hallazgos dentro de un amplio marco teórico.

Por su parte, Mischell (1968 en Fernandez Ballesteros, 1992) atacó frontalmente la teoría de los rasgos al subrayar el hecho de que lo que típicamente se observa es la inconsistencia de la conducta pese a que la teoría de los rasgos predice precisamente esta consistencia; sostuvo la determinación situacional de la conducta y apartir de ahí, ha ido creciendo la importancia que se otorga a la situacionalidad: las expresiones de los rasgos dependen de muchos elementos de la situación. Mischel se refiere a una inconsistencia en las respuestas sobre todo en el campo del temperamento. El valor de esta aportación radica en que reduce la atribución exagerada asignada a los rasgos para predecir la conducta y abre camino a la investigación sobre la interacción situación-rasgo.

Esta preocupación por la situacionalidad data de hace bastante tiempo pero sólo en épocas recientes se han desarrollado modelos formales para estudiarla. (Buss y Poley, 1979) Este concepto de la situacionalidad se retomará en el capítulo 3 en relación con el falseamiento de las respuestas.

1.1.3 EL ANALISIS FACTORIAL

Existen varios métodos de análisis factorial. Sin intentar describir el procedimiento matemático en si que es bastante complejo se intenta dar una explicación somera del método

Esta técnica fue desarrollada por Spearman en Inglaterra en 1927 para investigar acerca de las aptitudes. Como resultado de sus estudios, formuló la hipótesis de que existe un factor general de aptitud que llamó Factor G que explicaría las interrelaciones entre las mismas. Posteriormente, Thurstone (1938) en EEUU aplicó distintos métodos del análisis factorial al mismo campo de las aptitudes. A diferencia de lo encontrado por Spearman, los resultados lo llevaron a identificar nueve factores que denominó aptitudes mentales primarias. Independientemente de las conclusiones a las que llegó cada uno, el desarrollo de la técnica del análisis factorial en si permitió aplicarla al estudio de la personalidad que, antes de ellos, se enfrentaba a la limitante de la medición de variables y la comprobación de las hipótesis. (Catell y Kline, 1982) A partir de sus inicios, esta técnica ha sido ampliamente estudiada y perfeccionada lo que, aunado al avance de los cálculos computarizados, ha permitido utilizarla con amplitud.

El análisis factorial opera de la siguiente forma, según un resumen muy explicativo preparado por Hall y Lindsey (1984)

1. Se inicia con una gran cantidad de medidas significativas y cuantificable de la conducta que representen diferentes aspectos esta.
2. Se aplica la técnica factorial para descubrir los factores subyacentes que determinan o controlan el cambio en las variables externas. El propósito es identificar un pequeño número de factores básicos cuya acción explique la mayoría de las variaciones operadas en esa gran cantidad de medidas inicialmente obtenidas.

3. El análisis factorial estima también el factor de saturación o el grado en que cada uno de los factores identificados contribuye a determinar cada medida o conjunto de puntajes.

4. Se elaboran medios más eficaces para evaluar los factores.

5. Se asigna un nombre al factor y se interpreta su significado psicológico con base en la naturaleza de las medidas particulares que tienen una gran carga de dicho factor.

Aiken (1996) explica el método como sigue:

"El método del análisis factorial consiste en reducir la cantidad de variables a un grupo de medidas al tomar en cuenta la superposición (correlaciones) entre las distintas medidas. En el campo de las pruebas psicológicas, el problema consiste en encontrar unos cuantos factores sobresalientes que puedan explicar la mayor parte de la varianza en un grupo de calificaciones de pruebas diferentes.

Todos los procedimientos distintos para extraer estos factores a partir de las calificaciones en las pruebas se basan en el teorema que afirma que:

la varianza total observada de una prueba es igual a la suma de la varianza debida a los factores que la prueba tiene en común con otras pruebas, la varianza específica de la prueba en sí y la varianza que se produce por los errores de medición.

La parte de la varianza observada que se debe a factores comunes se llama comunalidad de una prueba y la parte que se debe a factores específicos se denomina especificidad.

Al sustituir estos elementos en la fórmula de la confiabilidad resulta la siguiente ecuación:

$$\text{confiabilidad} = \text{comunalidad} + \text{especificidad}.$$

El análisis factorial empieza por una matriz de las correlaciones entre las pruebas de *n* personas. Esta matriz es simétrica porque las correlaciones en una hilera determinada son idénticas a aquellas en la columna correspondiente. No hay entradas en la diagonal donde concuerdan las hileras y columnas para cada prueba. La cuestión de qué valores deben colocarse sobre esta diagonal depende del procedimiento de análisis factorial que se siga o la teoría que guíe al investigador.

En un tipo de procedimiento, el método central o centroide de Thurstone, los cálculos de la comunalidad de las pruebas se ubican sobre la diagonal de la matriz de correlación. Por otra parte, el método de eje principal requiere que las entradas de la diagonal sean 1.00. Sin detenerse en la cuestión de qué valores colocar sobre la diagonal, debe enfatizarse que la elección afecta tanto el número de factores que se extraen como los valores relativos que se obtienen (las cargas de los factores) de cada prueba sobre cada factor.

El resultado intermedio de un análisis factorial típico es una matriz de factores original, sin rotación. Aquí ya se observa que el número de variables se ha reducido a cierto número de factores comunes.

ROTACION DE FACTORES

A esta matriz de factores original se aplica un procedimiento que se conoce como rotación de factores a fin de simplificar la interpretación de los factores obtenidos por el aumento de la cantidad de cargas positivas altas y bajas en las columnas de la matriz de factores.

Dependiendo del método de rotación en particular es posible obtener ya sea factores no correlacionados: **rotación ortogonal**, o factores correlacionados: **rotación oblicua**.

Si los factores que se obtienen están relacionados entre sí, es posible volver a aplicar los mismos métodos analítico-factoriales a esas correlaciones y se obtienen los factores de segundo orden también llamados factores de orden superior. Son como "factores de factores"

INTERPRETACION DE FACTORES

Una vez completados todos los cálculos estadísticos que requiere el análisis factorial y la rotación, el investigador se enfrenta a la tarea de interpretar los factores identificados. Esto lo logra al inspeccionar el patrón de cargas altas y bajas de cada prueba sobre los diferentes factores: mientras más alta sea la carga será más importante el factor sobre una prueba dada. El cálculo de la comunalidad o la varianza de los factores comunes y de la especificidad, la varianza de los factores específicos, también puede ayudar en el proceso de interpretación de los factores.

Identificados los factores, es tarea del investigador tanto explicarlos como asignarles un nombre y aquí se inicia la teoría" (Aiken, 1996)

En ningún análisis factorial hay un método que *a priori* determine la situación de los factores (ejes) en relación con cada uno de los demás dentro del espacio factorial. Para solventar este problema, Thurstone (1947 en Catell y Kline, 1982) propuso que los ejes fueran rotados hasta alcanzar la estructura simple y ésta se define como la consecución de factores que tengan algunas saturaciones altas y muchas saturaciones bajas. En busca de la estructura simple, se utilizan rotaciones oblicuas más que ortogonales. Mediante procedimientos correctos de rotación oblicua, el análisis factorial puede producir factores invariantes que pueden ser considerados como determinantes o incluso como agentes causales...esto nos capacita para definir de forma operativa los constructos más importantes de la esfera de la personalidad. (Catell y Kline, 1982)

Eysenck (1953 en Catell y Kline, 1982) describe un factor como "una proposición condensada de relaciones (lineales) que generan entre sí un conjunto de variables". El considera tres aspectos útiles, su capacidad de ser descriptivos de una forma más económica al reducir el número de variables, su capacidad para sugerir hipótesis y su capacidad para probarlas o rechazarlas. Por su lado, Royce considera al factor como "un constructo definido operativamente por sus saturaciones factoriales" y da como sinónimos de factor los siguientes: dimensiones, determinantes, unidades funcionales, parámetros y categorías taxonómicas.

Para finalizar con esta explicación, se refieren las ocho condiciones recomendadas por Catell (1973 en Catell y Kline, 1982) para que un experimento analítico factorial defina los rasgos unitarios.

1. Elección estratégica de variables.
2. Muestreo de las personas.
3. Fijación del número de factores, basada en pruebas objetivas.
4. Determinación de las comunalidades.
5. Resolución por rotación oblicua única.
6. Prueba de la significación de la estructura simple.
7. Comprobación del grado de invarianza del modelo a lo largo de la investigación.
8. Comprobación de la invarianza de la estructura de orden superior.

Las primeras dos condiciones son de primordial importancia. La elección de variables determina de principio el valor de los hallazgos finales. Un adecuado muestreo de variables, tan amplio como sea posible, acorde con los objetivos de la investigación es requisito indispensable. El número de variables varía según la etapa: en la fase exploratoria es necesario tener tantas muestras de variables como sea posible. En la fase de comprobación de hipótesis se puede reducir el número de variables siempre que se incluyan algunas variables indicativas.

La segunda condición se refiere a la necesidad de una selección amplia de sujetos y que la muestra sea al azar. Se utilizan muestras homogéneas en los casos en que se tiene algún conocimiento teórico de las variables implicadas y se están probando hipótesis. Existen reglas para el tamaño y las características de las muestras.

Las otras seis condiciones para un análisis factorial científicamente válido son aspectos técnicos destinados a obtener la estructura simple.

LAS TEORIAS FACTORIALES.

1.2 LA TEORIA FACTORIAL DE HANS JURGEN EYSENCK

Nace en Alemania en 1916 de la cual sale en 1934. Obtiene un doctorado en Filosofía y Psicología en la Universidad de Londres. Dentro de su experiencia laboral, funge como investigador Jefe del Hospital de Emergencias de Mill Hill y conferencista del Instituto de Psiquiatría de la Universidad de Londres. Es también profesor visitante de la Universidad de Pennsylvania y de la Universidad de California en Berkeley. Asimismo, ha sido profesor de Psicología de la Universidad de Londres así como director de los departamentos de Psicología de los hospitales Maudsley y Bethelam en Londres.

Ha realizado extensas investigaciones en el campo de la personalidad y otros afines. Para el año 1969, había publicado 34 libros, como autor, coautor o colaborador, y 300 artículos. Sus principales libros fueron publicados de 1947 a 1970 pero en 1991 seguía publicando.

Eysenck (1991 en Fernandez-Ballesteros, 1992) también ha ejercido la crítica. Su trabajo se ha caracterizado por un marcado desdén de las preconcepciones existentes de la personalidad. Al referirse a las 22 teorías de la personalidad enlistadas por Hall y Lindsay, sostiene que sólo pueden salvarse las que cumplen los requisitos de la ciencia natural y que, aplicando este criterio, caerían las teorías de Freud, Adler, Jung, Binswanger, Horney, Sullivan, Fromm, Erickson y Maslow, entre otros.

Este perseverante, polémico y prolífico investigador contribuyó a renovar el ímpetu de la investigación de la personalidad insinuando en la necesidad de aplicar los métodos de las ciencias naturales a estos estudios. Ha preferido concepciones simples y operacionales partiendo del principio de que la medida es fundamental en todo avance científico. Critica duramente las teorías de personalidad recargadas de variables complejas y no definidas. (Cueli y Reidl, 1972)

Antecedentes

Las teorías de Jung y Mc Dougall guiaron a Eysenck a seleccionar las 39 escalas que parecían relevantes en una muestra anormal con el fin de encontrar factores de extroversión y neuroticismo. Jung en su obra "Tipos psicológicos" (1923) describió al introvertido y al extrovertido. Eysenck consideró también estudios factoriales preliminares realizados por Spearman aún cuando la metodología que éste usó era muy sencilla. Asimismo, tomó en cuenta las investigaciones de Heyman y Wiersma los cuales habían identificado tres dimensiones: emocionalidad, actividad y predominio de la función primaria o secundaria.

También los hallazgos de Webb de 1915 le sirvieron de antecedente. Eysenck revisó esta investigación en 1970 y encontró un factor llamado W el cual es contrario a la emocionalidad. También la teoría de Kretschmer ha servido de antecedente y de base para generar investigación y confirmar o rechazar sus hallazgos. Los factores identificados inicialmente por Guilford fueron la base para los primeros estudios de Eysenck sobre la extroversión. (Catell y Kline, 1982)

Su teoría de la personalidad tiene una primera meta, la clasificación de sujetos por sus características y la utilización del análisis factorial en la formación de las taxonomías. Busca sobre todo, identificar dimensiones primarias de la personalidad para lograr una poderosa tipología única.

Su definición de personalidad de 1947 parece particularmente acorde con la de Allport. Define la personalidad como "la suma total de los patrones de conducta actuales o potenciales determinados por la herencia y el medio que se originan y desarrollan en una interacción funcional de los cuatro sectores principales dentro de los cuales están organizados estos patrones de conducta: el sector cognoscitivo, (la inteligencia), el sector conativo (el carácter), el sector afectivo (el temperamento) y el sector somático (la constitución)" (Cueli y Reidl, 1972) Esta definición evidencia su cercanía con las teorías de las constituciones y temperamentos pero el hecho de utilizar la metodología factorial le confiere una distinción teórica y metodológica significativa (Fernandez-Ballesteros, 1992); incluso, ha realizado bastantes investigaciones utilizando la estructura corporal de un individuo como variable de personalidad.

Estructura de la personalidad

Su modelo de personalidad se caracteriza por un pequeño número de dimensiones principales definidas con gran cuidado.

Eysenck considera que la personalidad esta compuesta por actos y disposiciones organizada como una jerarquía por su generalidad e importancia en cuyo primer nivel se encuentran los tipos, identificados como factores generales; en el nivel siguiente están los rasgos que corresponden a los factores de grupo, posteriormente se ubicarían las respuestas habituales que serían los factores específicos y, las respuestas específicas, las que se presentan eventualmente, serían el factor error.

Si se repiten las respuestas específicas se convierten en respuestas habituales; algunas de ellas se interrelacionan y tienden a existir juntas en una persona formando lo que Eysenck llamó un rasgo y éstos a su vez se organizan para formar los tipos. Insiste en que los rasgos deben definirse operacionalmente (Cueli y Reidl, 1972)

Dimensiones de la personalidad. Eysenck analiza la personalidad a lo largo de tres dimensiones amplias:

RASGOS DE PERSONALIDAD DE EYSENCK

neuroticismo vs. no-neuroticismo
extraversión vs. introversión
psicoticismo vs. no-psicoticismo

Inicialmente identificó solamente los dos primeros factores y a ellos les ha dedicado gran parte de sus investigaciones y posteriormente consideró el tercer factor, el psicoticismo. Admite que estas tres dimensiones no son necesariamente las únicas alternativas.

Existe un esquema famoso de Eysenck y Rachaman que considera las dos dimensiones básicas, extroversión-introversión y estabilidad-inestabilidad o neuroticismo, los 32 rasgos que corresponden a cada dimensión y su correspondencia con la antigua clasificación de tipos de personalidad de Hipócrates (Mischel, 1988).

La concepción de la estructura de la personalidad de Eysenck destaca por su simplicidad. Reduce el número de tipos y los integra a una teoría biológica que explica su distinción y sus características. Plantea que la personalidad tiene bases biológicas que radican en el sistema nervioso del hombre, en particular, en relación a los procesos de excitación e inhibición corticales. La diferencia se basa en la calidad del sistema nervioso en orden a la extensión y duración de su activación, siguiendo la tradición pavloviana. También otorga a los factores P, E y N un alto grado de heredabilidad. La distinta activación combinada del SNA y el SNC da lugar a la aparición de los tipos (Fernández-Ballesteros, 1992).

Eysenck y sus colaboradores han utilizado tests de clasificación, cuestionarios, tests situacionales y otras medidas psicológicas para identificar factores o ampliar las investigaciones sobre sus hallazgos. A pesar de las diferencias y revisiones de estos instrumentos se mantiene una fuerte continuidad entre ellos.

Eysenck ha recopilado una gran parte de la literatura que muestra cómo sus instrumentos, usados por otros investigadores, conservan sus cualidades psicométricas y la misma estructura factorial lo que resulta en una validación de los mismos. (Fernández-Ballesteros, 1992)

Una lista cronológica de los principales inventarios de Eysenck se presenta a continuación:

FECHA	NOMBRE	SIGLAS	FACTOR EVALUADO
1952	Cuestionario Medico Maudsley	MMQ	neuroticismo
1959	inventario personalidad Maudsley	MPI	neuroticismo, extraversión y escala de mentira
1964	Inventario de personalidad de Eysenck	EPI PEN	igual que el MPI extrov. neuroticismo y psicoticismo
1976	Cuestionario de Personalidad de Eysenck-EPQ versiones para jóvenes y adultos	EPQ	extrov. neuroticismo y psicoticismo

(Fuente: Cateil y Klim, 1982)

Del Inventario de personalidad de Eysenck, EPI, existe una adaptación española de 1973. Del Cuestionario de personalidad de Eysenck, EPQ existe una versión española estandarizada (Eysenck y Seisdedos, 1978, Eysenck et al, 1982 y Perez et al. 1990) Existe también una adaptación mexicana de Lara et al de 1992. (en Fernández-Ballesteros, 1992).

Metodología.

Aborda la investigación con el método hipotético-deductivo. Utiliza métodos estadísticos multivariados sobre todo análisis factoriales. Considera este investigador que las dimensiones deben identificarse antes de elaborar una teoría de la personalidad. Insiste en el uso de definiciones operacionales, es decir, que debe explicarse el procedimiento para llegar a los resultados.

Ha desarrollado diferentes métodos como la técnica del análisis de criterio. Esta es una innovación metodológica importante dentro del análisis factorial y consiste en adecuar un factor en el análisis para lograr una separación máxima para un grupo de criterio determinado. Consiste en contar con dos muestras de sujetos que estén contrastados con el grado en que posean una variable en particular y aplicar una prueba o serie de ellas para determinar qué tan sensible es la prueba o las pruebas a dicha variable. En otras palabras, los grupos criterios implican que el último factor derivado se inclinara más excesivamente sobre esas pruebas que discriminen más efectivamente sobre los grupos criterio. Este método lo aplicó en estudios del factor neuroticismo realizados con grupos de soldados normales y neuróticos. (Cuelli y Reidl, 1972.)

Eysenck se ha preocupado por reunir datos de todas las investigaciones afines que confirmen su teoría. Este autor ha destacado por la amplitud de sus investigaciones, por su rigor metodológico, por su cautela al enunciar teorías, por el orden y energía al dedicarse a temas de su interés, que es muy amplio y variado, y por su capacidad para generar y aplicar métodos en donde no se habían utilizado.

1.3 LA TEORIA FACTORIALISTA DE JOY PAUL GUILFORD (1897-1987)

Nace en EE.U.U. y se gradúa con honores en el Colegio de Aurora. Ingres a la Universidad de Nebraska, estudia Química y posteriormente Psicología mientras colabora como asistente de laboratorio y de enseñanza. Obtiene un doctorado y trabaja en la Universidad de Illinois. De 1928 a 1940 colabora con la Universidad de Nebraska y posteriormente en la Universidad de California del Sur. Fué presidente de la *American Psychological Association* en la década de los cincuenta. Publico 25 libros, de 1950 a 1967, cerca de 300 artículos y desarrolló 30 pruebas psicométricas. Sus principales libros, *Psychometric Methods* (1954), *Personality* (1959), *The Nature of human Intelligence* (1967) y *Fundamental statistics in Psychology and Education* (1965)

Esta breve nota biográfica da apenas una leve idea de una vida dedicada al estudio y a la investigación en Psicología. Abarco campos de estudio de la capacidad intelectual, intereses, creatividad, personalidad, etc. Aplico primordialmente el análisis factorial en sus investigaciones.

Se puede considerar a Jung como un antecedente importante de este trabajo pues en 1934 Guilford intentaba estudiar la extroversión elaborando elementos que fueran representativos de la típica descripción junguiana.

Estructura de la personalidad

Define el rasgo como "cualquier aspecto distintivo y duradero en el que un individuo difiere de otros". Los rasgos pueden ser somáticos (fisiológicos y morfológicos), actitudinales, temperamentales y motivacionales (intereses, aptitudes y necesidades). Los rasgos deben ser factibles de medir o escalar, son consistentes en su posición relativa entre unos y otros, son durables y flexibles, universales, su generalidad puede ser mayor o menor en una persona, se clasifican en rasgos fuertes y débiles, se organizan en patrones; algunos rasgos se relacionan entre si y otros son independientes. (o rasgos primarios) (Cueli y Reidl, 1972)

Guilford concuerda con Eysenck en la clasificación de niveles jerárquicos de los rasgos, de los tipos altos superiores, los rasgos primarios, los hexes (disposiciones más o menos específicas como los hábitos) y en el nivel inferior, las acciones específicas.

Reconoce asimismo las siguientes subregiones de la personalidad: dimensiones de capacidad, dimensiones de temperamento y dimensiones hórnicas, que corresponden más o menos a los rasgos de capacidad, de temperamento y dinámicos de Cattell. Agrega las dimensiones somáticas y las dimensiones patológicas para incluir los trastornos de personalidad. Organiza y vincula las dimensiones de cualquiera de estas subregiones en tablas o matrices de rectángulos bi o tridimensionales en donde un factor determinado se ve como una función o cualidad más general y se expresa en la conducta. Por ejemplo, en el ámbito del temperamento, una dimensión "positiva-negativa" produce en la conducta general un factor de "seguridad-inferioridad" y en el ámbito emocional un factor de "jovialidad-timidez". Estas ligas no se descubren solamente con el análisis factorial sino representan interpretaciones o esquemas en donde pueden ordenarse los factores ya descubiertos como en su modelo de Estructura del intelecto. (Hall y Lindzey, 1984)

Clasificación de las dimensiones identificadas por Guilford.

Dimensiones somáticas

Realizo 32 análisis factoriales utilizando medidas morfológicas e identifico seis rasgos corporales como el largo del cuerpo y el grosor muscular. También identifico cuatro rasgos fisiológicos: a) dominio simpático contra parasimpático b) tensión c) funcionamiento tiroideo d) nivel de glucosa en sangre.

Dimensiones de aptitud

Considera la aptitud como la dimensión subyacente de la habilidad. Las habilidades primarias caen dentro de tres campos: perceptual, psicomotora e intelectual. Dentro de estas últimas, se distinguen las habilidades de la memoria y las del pensamiento. Las del pensamiento a su vez, se dividen en habilidades cognitivas, productivas y evaluativas. A partir de lo anterior, Guilford elaboró un modelo muy elaborado: un esquema tridimensional denominado "La Estructura del Intelecto" (1956).

Dimensiones del temperamento

Los rasgos temperamentales tienen que ver con la manera en que ocurren las acciones de un sujeto. Los clasifica como generales, emocionales y sociales y en general son rasgos bipolares.

Dimensiones motivacionales de la personalidad

No es posible separar completamente las dimensiones intelectuales o temperamentales de las de la motivación. Bajo este rubro se encuentran los *drives* o pulsiones biológicas que dependen de las necesidades orgánicas. Dentro de estas dimensiones se incluyen las necesidades, los intereses y las actitudes.

Dimensiones de la patología

Su interés en identificar este tipo de dimensiones surgió de la confusión que percibía en las categorías nosológicas tradicionales. Considero que un enfoque más sano debería partir del estudio de la concomitancia de los síntomas que significaría realizar estudios correlacionales de los síntomas. Aplicando el análisis factorial se obtuvieron varias dimensiones de neurosis y de psicosis verificadas en sus diferentes grados de complejidad patológica.

(Fuente: Cueli y Reidt, 1972)

Instrumentos de Guilford para las investigaciones sobre rasgos de personalidad.

Guilford construyó el el Inventario de Factores, STDCR (1940) y el Inventario de Factores Guilford-Martin, GAMIN (1943) que sirvieron de base para sus primeros estudios. Posteriormente, ambos fueron combinados y dieron lugar al Inventario de personalidad llamado Examen de temperamento Guilford-Zimmerman, GZTS, en 1949. Este mide diez rasgos definidos factorialmente: actividad general, apocado y ratimia (disposición a la alegría), dominio, sociabilidad, estabilidad emocional, objetividad, amistad, solicitud, relaciones personales y masculinidad; sin embargo, Eysenck considera que la descripción de 1956 es la más lograda:

Principales factores de personalidad (temperamento) de Guilford

G Actividad General	D Depresión
A Ascendencia	C1 Emocionalidad
M Masculinidad-feminidad	R Restricción vs ratimia
I Confianza vs. sentimientos de inferioridad	O Objetividad
N Calma y compostura vs. nerviosismo	Ag Agradable (vs. agresivo-hostil)
S Sociabilidad	Co Cooperación y tolerancia
T Reflexividad	

(Fuente: Cattell y Kline, 1982; Hall y Lindzey, 1984)

Metodología.

Guilford es de los primeros investigadores del análisis factorial. Es conocido por sus textos de estadística y métodos psicométricos. Inició su estudio acerca de los rasgos en los años treinta influido por Thurstone y su obra "Los vectores de la mente". Prefiere el modelo multifactorial sobre el modelo de Spearman de factor g . En sus análisis factoriales, aplica el método de rotación ortogonal a diferencia, por ejemplo, de Cattell que utiliza el método de rotación oblicua.

La información de la que parte son cuestionarios de autodescripción y datos de la observación de la conducta por otros.

Cattell hace una crítica del aspecto técnico de los análisis factoriales de Guilford y cuestiona la precisión de sus hallazgos pero por otra parte, considera que podría haber proseguido sus investigaciones sobre la misma línea. Deplora que no haya relacionado estos factores con una amplia variedad de otras variables psicológicas y que no haya considerado los factores de segundo orden. (Cattell y Kline, 1982)

Las conclusiones de Guilford en ocasiones adolecen de una complejidad que las convierte en difíciles de comprender.

1.4 LA TEORIA FACTORIALISTA DE RAYMOND B. CATELL.

Nació en 1905 en Staffordshire, Inglaterra. Obtuvo el título de *Bachelor* y *Master* en ciencias y en 1929 el de Doctor en Filosofía por la Universidad de Londres. Fue conferencista en Exeter, catedrático universitario y director de una clínica de Psicología en Leicester, Inglaterra. Obtuvo el doctorado en ciencias por la Universidad de Londres. Posteriormente fué investigador en la Universidad de Columbia. Fue asimismo investigador en la Universidad de Illinois. Ha sido miembro de importantes asociaciones de psicólogos y fundador de la Sociedad de Psicología experimental multivariada.

Los principales libros son *The description and measurement of personality* (1946), *Personality: a systematic, theoretical and factual study* (1950), *Personality and motivation structure and measurement* (1957), *The scientific analysis of personality* (1965), *Handbook of multivariate experimental psychology* (1966) y *Handbook of modern personality theory* (1971).

Ha publicado un gran número de tests. Los principales son: Los tests de inteligencia libres de cultura. La batería de tests analítico-objetivos. El cuestionario de los dieciséis factores de personalidad 16 PF. El cuestionario para adolescentes-bachilleres HSPQ, pruebas para niños CPQ y ESPQ. MAT de motivaciones. SMAT, el cuestionario de análisis clínico CAQ. La batería de los Ocho Estados (de ánimo), la Escala de Ansiedad y muchos más.

Su obra ha sido amplísima como autor o coautor de libros, artículos, manuales o compilaciones. En 1964 había publicado 235 artículos y 22 libros él solo y 22 en colaboración con otros investigadores.

Antecedentes del trabajo de Cattell

Un autor importante es Gordon Allport y sus trabajos referentes a los rasgos de la personalidad. Cattell retomó una lista de palabras descriptivas de la personalidad elaborada por Allport y Odbert que inicialmente ascendía a casi 18,000 adjetivos y ésta fué la base de sus estudios sobre los rasgos de personalidad. Cattell, como otros investigadores, consideraba que existe una concordancia entre el lenguaje y la esfera total de la personalidad: un rasgo existe en la medida que existe el término para designarlo. Por otro lado, la clasificación de Cattell de los rasgos tiene similitud con la de Allport.

Sigue los pasos de Kurt Lewin en cuanto a trasladar las ideas psicológicas a formas matemáticas explícitas. Quizá sea Henry Murray el teórico más cercano a Cattell. Ambos tienen un concepto amplio de la personalidad y ambos han desarrollado sistemas teóricos inclusivos. Se interesaron por el aspecto empírico del ámbito de la personalidad y elaboraron gran cantidad de constructos ligados funcionalmente con los datos. Ambos estudiaron a profundidad la motivación. Además, se han caracterizado por su gusto por los neologismos. Una diferencia fundamental es el empleo del análisis factorial por Cattell. Sus formulaciones teóricas están relacionadas también con McDougall que se interesó en las dimensiones subyacentes de la conducta. Freud y los autores consagrados al psicoanálisis proporcionaron material para sus investigaciones sobre el desarrollo y muchas de las ideas teóricas de Cattell a este respecto guardan estrecha vinculación con Freud (Hall y Lindzey, 1984)

Definición de personalidad.

Catell utiliza la palabra personalidad para referirse a la totalidad de la conducta. Define la personalidad como "aquello que permite predecir lo que hará una persona en una situación dada". La personalidad se refiere y se deduce de todas las relaciones humanas entre el organismo y su medio e incluye tanto conductas manifiestas como encubiertas. Catell considera a la personalidad como una estructura compleja y diferenciada de rasgos cuya motivación depende en gran medida de un subconjunto llamado rasgos dinámicos. (Hall y Lindzey, 1984 y Mischel, 1988)

El rasgo como enfoque estructural básico.

Dos aspectos son fundamentales en el trabajo de Catell, por un lado el enfoque de rasgos en el estudio de la personalidad y por otro lado, el método factorial utilizado para abordarlo. Los rasgos son tendencias relativamente estables a comportarse de determinada manera. Rasgo es una "estructura mental inferida a partir de la conducta destinada a explicar la regularidad o uniformidad de esta" (Mischel, 1988) Cualquier característica física, fisiológica, psicológica o sociológica puede ser considerada un rasgo.

CLASIFICACION DE RASGOS DE CATELL

RASGOS COMUNES. Aplicables a todos o casi todos los individuos.

RASGOS UNICOS. Aplicables a un solo individuo.

RASGOS FUENTE O FUNDAMENTALES.(FACTORES) Correlaciones observadas entre diversas variables. Operan como una fuente subyacente o causa de la conducta observada. Sólo se identifican mediante análisis factorial.

Estos se dividen en dos: los que reflejan condiciones ambientales y los que reflejan herencia. **RASGOS SUPERFICIALES. (AGRUPAMIENTOS)** Son conjuntos de elementos, respuestas, de rasgos declarados o manifiestos que parecen ir juntos.

RASGOS RELATIVAMENTE ESTABLES

ESTADOS FLUCTUANTES (Equivalentes a estados de ánimo)

De acuerdo con la modalidad por la cual se expresan:

RASGOS DE HABILIDADES. Los relacionados con la eficiencia en logro de objetivos

RASGOS TEMPERAMENTALES. Los asociados con la energía o reactividad emocional.

RASGOS DINAMICOS. Tienen que ver con la motivación o los intereses. Los rasgos relevantes para que un individuo se ponga en acción respecto de un objetivo.

Catell considera tres formas de obtener datos: **Datos L (Life)** Se obtienen por valoración directa de los sujetos. Son calificaciones de la vida real, informes de la conducta de una persona en sociedad o evaluaciones realizadas por otros.

Datos Q (Questionnaire) Se obtienen mediante autoinformes con entrevistas, escalas, inventarios. Se basan en la introspección. La apreciación personal de la conducta.

Datos T (Test) Se obtienen de la respuesta que un sujeto desarrolla ante unas determinadas situaciones de prueba. Se les denomina pruebas objetivas debido a que el sujeto reacciona sin saber cuál de los aspectos de su conducta se está evaluando.

El análisis factorial depende de los datos que sirvan de base y Catell puso especial énfasis en la recolección de datos. De los datos L y Q emergen rasgos aptitudinales o temperamentales. De los datos T surgen datos motivacionales o dinámicos. Los datos Q son adecuados para evaluar rasgos de temperamento pero resultan inconvenientes para evaluar habilidades y cuestionables para evaluar motivaciones. (Cuelli y Reidl, 1972 y Fernandez-Ballesteros, 1992)

Se considera que si los rasgos obtenidos en los cuestionarios o inventarios son reales en algún sentido, es decir, válidos, deben corresponder a rasgos derivados por otros métodos y Catell descubrió una gran coincidencia entre los rasgos de los datos L y de los datos Q pero no la encontró así con los rasgos de los datos T probablemente porque esa técnica no se ha desarrollado lo suficiente. El tomar en cuenta tanto los datos L como los datos Q respalda la pretensión de Catell de abarcar completamente los rasgos originales de la personalidad (Buss y Poley, 1979)

La esfera de la personalidad del lenguaje

Catell ha sostenido que el lenguaje es la única fuente efectiva de la totalidad de los rasgos de la personalidad. Considera que todos los aspectos discernibles del comportamiento deben haberse verbalizado en el transcurso del tiempo en idiomas con una importante tradición literaria como es el inglés y que la esfera de la personalidad del lenguaje proporciona un buen cuadro de la totalidad de la personalidad.

Tomó como base de sus estudios la lista de casi 18,000 términos del idioma inglés elaborada por Allport y Odbert Utilizando diversos procedimientos la redujo a 4,500 adjetivos y posteriormente a 171 terminos. Estos fueron analizados factorialmente y se obtuvieron 36 conglomerados de correlaciones -rasgos superficiales- a los cuales se añadieron diez rasgos más. Estudios posteriores redujeron este número a 16 factores que constituye la base de su test más conocido. Este proceso de destilación representa la "única contribución trascendental al muestreo extenso del campo de la personalidad y el descubrimiento del número de los rasgos originales que se requieren para su justificación" (Catell y Kline, 1982; Buss y Poley, 1979)

LOS RASGOS DE PERSONALIDAD

La tabla siguiente muestra los rasgos primarios normales identificados en las múltiples investigaciones de Catell. Estas dimensiones están basadas en análisis factoriales realizados a lo largo de los años, en ocho estudios que involucraron a varios miles de sujetos. En el octavo estudio, por ejemplo, el número ascendió a 6,500 personas. En algunos casos, se usaron intercorrelaciones entre elementos; en otros, correlaciones entre paquetes de elementos. Esta descripción está tomada de Catell, 1973, (en Catell y Kline, 1982) Las pruebas usadas fueron el 16 PF, para adultos, el HSPQ para adolescentes bachilleres y el CPQ para niños de escuela primaria.

Los factores están en orden decreciente, de acuerdo con su importancia. Expuesto de manera más precisa, son 14 rasgos de temperamento, uno de aptitud, el B, y uno de motivación, el Q4. Como se ha mencionado, Catell y sus colaboradores definieron la personalidad en términos de palabras descriptivas del diccionario a partir de una muestra de adjetivos realmente extensa, clasificaron a un grupo de personas por su comportamiento habitual (datos L) en esos rasgos y sometieron las correlaciones resultantes a un análisis factorial. Por otro lado, diseñaron cuestionarios y pruebas objetivas (datos Q y datos T), las aplicaron a diferentes muestras de sujetos y analizaron factorialmente los resultados procurando alcanzar la estructura simple. Los factores resultantes han sido sometidos a estudios posteriores para facilitar su identificación. (Catell y Kline, 1982) y la tabla muestra precisamente los 16 rasgos primarios.

Una vez que un rasgo o factor es identificado, se le asigna un nombre y se interpreta su significado psicológico, también usando el lenguaje como herramienta. Catell utilizó muchos neologismos al designar factores, por ejemplo, el factor A- llamado Sizia o Sizothimia proviene de los términos latinos sizo=plano y lo uso para referirse a la emotividad seca y plana en el individuo.

(Catell, Eber y Tatsuoka, 1980) Los términos explicatorios varían de fuente a fuente e incluso de país a país del mismo idioma, por ejemplo, en el manual del Cuestionario del 16 PF, traducida en México y en la obra de Karson y O'Dell (1989), traducida en España.

Catell asignó las letras del alfabeto a los rasgos primarios pues su modelo consistentemente consideraba los 16 factores primarios que fueron identificados inicialmente con propiedad. En investigaciones posteriores identificó los "siete rasgos ausentes" a los que les corresponden las letras D, J, K, P, Q5, Q6 y Q7 e incluso desarrollo, con sus colaboradores, un complemento de la prueba 16 PF. Los factores denominados Q con algún subíndice indican rasgos encontrados en datos Q.

Catell consideró aproximadamente 16 rasgos primarios y ocho secundarios; este número variaba con los sucesivos estudios y con frecuencia habla de 22 rasgos que resultan de las suma de los quince primarios, eliminando el B y los siete ausentes o habla de 23 rasgos, incluyendo B.

DESCRIPCIÓN DE LOS RASGOS PRIMARIOS NORMALES.

Indice Rasgo fundamental	puntuaciones bajas	puntuaciones altas
A	SIZIA Reservado, alejado, crítico, solitario, rígido	AFFECTIA Abierto, afectuoso, participativo, inalterable
B	INTELIGENCIA BAJA Torpe	INTELIGENCIA ALTA Listo
C	POCA FUERZA DEL EGO Afectado por los sentimientos, emocionalmente inestable, poco estable, fácilmente perturbable	MUCHA FUERZA DEL EGO Emocionalmente estable, maduro, afronta la realidad, tranquilo
E	SUMISO Sumiso, apacible, manejable, dócil, acomodaticio	DOMINANCIA Asertivo, agresivo, competitivo, obstinado
F	DESURGENCIA Sobrio, factuario, serio	SURGENCIA Descuidado, alegre, entusiasta
G	POCA FUERZA DEL SUPEREGO Despreocupado, le tienen sin cuidado las normas.	MUCHA FUERZA DEL SUPEREGO Conciencioso, metódico, moralista, formal
H	TRECTIA Cohibido, tímido, susceptible	PARMIA Emprendedor, desahogado, socialmente atrevido
I	HARRIA Sensibilidad dura, confiado en sí mismo	PREMSIA Sensibilidad blanda, sensible, superprotegido, pendiente
L	ALAXIA Confiable, adaptable a las circunstancias	PROTENSION Suspica, difícil de engañar
M	PRAXERNIA Práctico, con los pies en la tierra, comprometido	AUTIA Imaginativo, bohemio, abstracto
N	SENCILLEZ Franco, natural, auténtico pero poco burdo en lo social	ASTUCIA Astuto, calculador, hábil en lo social
O	ADECUACION IMPERTURBABLE Aprensivo, con remordimientos, inseguro, preocupado, inquieto	TENDENCIA A CULPABILIDAD Seguro de sí, apacible, tranquilo, satisfecho y sereno
Q1	CONSERVADURISMO Conservador, de ideas tradicionales arraigadas	RADICALISMO Liberal, experimental, librepensador
Q2	ADHESION AL GRUPO Dependiente del grupo, seguidor nato, se apunta a todo	AUTOSUFICIENCIA Autosuficiente, lleno de recursos, prefiere sus propias decisiones
Q3	BAJA INTEGRACION DE SUS PROPIOS SENTIMIENTOS Indisciplinado, autoconflictivo, sigue sus propios impulsos, no le atan las normas sociales	MUCHA FUERZA DE SUS PROPIOS SENTIMIENTOS Controlado, estanco en cuanto al poder de su voluntad, cuidadoso en lo social, compulsivo, atento de su propia imagen
Q4	POCA TENSION ERGICA Retajado, tranquilo, alejargado, sosegado, no frustrado	MUCHA TENSION ERGICA Tenso, frustrado, agobiado en exceso

Catell y Kline, 1982

LOS FACTORES DE SEGUNDO ORDEN DE CATELL

Estos son derivados de las factorizaciones de orden superior, es decir, factores derivados de las correlaciones entre los factores de primer orden. No son factores secundarios por su importancia sino por el método seguido para obtenerlos. Estos se obtienen a partir de datos Q. Se basan en catorce estudios factoriales de adultos con el 16 PF. Cattell y Nichols (1972) han hecho especiales esfuerzos para casar los factores de orden superior que procedían de distintos estudios (en Cattell y Kline, 1982). Seis de estos factores han sido verificados exhaustivamente en España en 1980 por Seisdedos y Prieto. (Seisdedos, 1985)

FACTORES DE SEGUNDO ORDEN EN ADULTOS NORMALES

Nombre	Descripción en términos de factores primarios
1. Exivia	O sea, extraversión Sociable (A), surgente (F), emprendedor (H), dependiente (Q ₂)
2. Ansiedad	Poca fuerza del ego (C), tímido (H), suspicaz (L), propenso a la culpabilidad (O), sentimiento propio bajo (Q ₃), tenso (Q ₄)
3. Corterba	Insociable (A-), insensible (I-), astuto (N)
4. Independencia	Surgente (F) dominante (E), emprendedor (H), desentendido (M), suspicaz (L)
5. Discreción	Astuto (N), sociable (A)
6. Subjetividad	Desentendido (M), radical (Q ₁)
7. Inteligencia	Inteligencia (B)
8. Buena edificación	Supremo (G), sumiso (E-), desurgente (F), sentimiento propio (Q ₃) (Cattell y Kline, 1982)

LA ECUACION DE ESPECIFICACION.

Cattell buscaba la manera de reunir los datos de capacidad, temperamento, etc. para predecir la respuesta de un individuo en una situación determinada; propone una fórmula matemática para lograrlo: la ecuación de especificación. Esta es la representación multidimensional tanto de la persona como de la situación psicológica.

El modelo incluye la integración de rasgos, estados anímicos y situaciones mediante la asignación de pesos distintos que modifican las calificaciones de los rasgos. La situación global se conceptualiza por dos componentes, el focal y el ambiental o contextual de situación. Esto permite teóricamente una predicción más precisa de la conducta que aquella derivada únicamente de los rasgos.

$$\text{ECUACION DE ESPECIFICACION}$$

$$R = s_1 T_1 + s_2 T_2 + s_3 T_3 + \dots + s_n T_n$$

La respuesta puede predecirse por las características de una persona determinada (los rasgos T1 a Tn) ponderados según su importancia en la situación real (los índices situacionales s1 a sn).

Si un rasgo determinado es muy importante para una respuesta dada, la s correspondiente será amplia; si el rasgo es insignificante, la s será cero; si el rasgo se opone o inhibe la respuesta, el signo de S será negativo. La fórmula de la ecuación implica que cada rasgo tiene un efecto independiente y adicional sobre la respuesta. (Hall y Lindzey, 1984)

LOS RASGOS DINAMICOS

Los supuestos que sirvieron de base a Catell para la investigación en el campo de la motivación fueron las siguientes:

El hombre hereda ciertas pulsiones básicas que proveen los resortes principales originales para la acción.

Ciertos conjuntos de hábitos y actitudes se aprenden de tal forma que es posible reconocer un número de estructuras unitarias integradas.

Un segmento sustancial de la motivación es inconsciente.

Un sentimiento de sí mismo surge por aprendizaje y va logrando control sobre las pulsiones básicas, transformándolas en expresiones legítimas y socialmente aceptables.

El conflicto puede surgir entre diferentes estructuras dinámicas, entre el yo y el ello, entre el yo y el superyo.

La mayoría de la conducta clínica neurótica se origina en desbalances entre los poderes ejercidos a través de estas estructuras y de los conflictos resultantes.

Catell (1957 en Cueli y Reidl, 1972) considera que toda conducta es motivada y que un motivo es algo que existe en el aquí y en el ahora al actuar.

CLASIFICACION DE RASGOS DINAMICOS

Los rasgos dinámicos superficiales son de tres clases: ergs o ergios, actitudes y sentimientos. Los ergs corresponden aproximadamente a los impulsos biológicos; los sentimientos a las estructuras de actitud adquiridas.

ERGS

Los ergs son rasgos originarios, constitucionales, dinámicos, que así concebidos permiten a Catell representar los móviles de conducta congenitamente determinados aunque modificables. Los ergs son las fuerzas impulsoras básicas de la personalidad. El énfasis en la motivación ergica refleja su convicción de que los psicólogos norteamericanos han subestimado la importancia de las determinantes hereditarias de la conducta.

Su definición de erg incluye cuatro puntos: la respuesta perceptual, la respuesta emocional, los actos instrumentales que conducen al objetivo y la satisfacción final en sí misma. Los ergs son disposiciones psicofísicas innatas que permiten al individuo adquirir reactividad (atención, reconocimiento, respuesta) ante ciertos estímulos más rápidamente que ante otros. Los ergs se pueden considerar pulsiones pues reúnen las características de éstas de acuerdo con McDougall: proclividad espontánea a atender ciertos objetos y situaciones otorgándoles mayor importancia relativa; acompañados de una emoción característica que es bastante específica la pulsión y su acción y poseen al final una meta particular.

Mediante análisis factorial, Catell y sus colaboradores han identificado siete ergs: sexo, sociabilidad, protección parental, curiosidad, huida (miedo), autoafirmación y sexualidad narcisista (relacionado con la autoindulgencia general)

ACTITUDES

Una actitud es la variable dinámica manifiesta, la expresión observada de una estructura dinámica subyacente de la cual deben inferirse los ergs y los sentimientos y sus interrelaciones. La actitud es una estructura que se infiere de la aparición repetida de respuestas conductuales de cierto tipo. Catell y sus colaboradores han interrelacionado de 60 a 70 diferentes instrumentos para medir la fuerza de la actitud con sus componentes conscientes e inconscientes. Se identificaron cinco factores componentes de la actitud denominados de alfa hasta épsilon y se relacionaron especulativamente con lo conceptos psicoanalíticos del ello, el yo, los complejos reprimidos, etc. (Hall y Lindzey, 1984; Cueli y Reidl, 1972)

CONFIGURACION DE LOS FACTORES DE LA MOTIVACION

Nombre	Descripción
1 Alpha	Ello consciente
2 Beta	Interés vivenciado
3 Gamma	Superego
4 Delta	Ello fisiológico
5 Epsilon	Conflicto inconsciente
6 Zeta	No identificado
7 Eta	No identificado

(Fuente: Catell y Kline, 1982)

Estos factores representan hallazgos experimentales comprobables, de una precisión poco usual en los conceptos y variables de las teorías clínicas.

SENTIMIENTOS

Un sentimiento es un rasgo de fuerza dinámica, moldeado por el ambiente, resultado de factores vivenciales o socioculturales. El sentimiento constituye una fuente de energía e interés existente pero solo acompaña a la energía derivada básicamente de los ergos. Los sentimientos son una estructura dinámica como los ergs pero, a diferencia de éstos, se pueden desintegrar y aún desaparecer.

Algunos de los sentimientos encontrados por Catell y sus colaboradores en investigaciones con adultos jóvenes, principalmente del sexo masculino, son: carrera o profesión, deportes y juegos, intereses rutinarios, religión, padres, esposa o novia y si mismo.

El sentimiento de si mismo es uno de los más estables y coherentes de los sentimientos estudiados. Se vincula con la expresión de la mayoría de los ergs. Catell considera que este sentimiento juega un papel decisivo en la integración de la personalidad. En el análisis de este sentimiento se puede observar la influencia de los conceptos psicoanalíticos del yo y del superego, del yo de Allport y del sentimiento del propio interés de McDougall. (Hall y Lindzey, 1984; Cueli y Reidl, 1972)

La trama dinámica.

Los diversos rasgos dinámicos se interrelacionan en una pauta de *subsidio*, esto es, que ciertos elementos son subsidiarios de otros o sirven como medios para sus fines. En general, las actitudes son subsidiarias de los sentimientos y estos de los ergs. Estas intrincadas relaciones se pueden representar gráficamente en una red o "trama dinámica" como parte de la estructura motivacional de un individuo hipotético. Lo que Catell hace es medir una cantidad de actitudes, analizarlos factorialmente e interpretar los factores como representantes de ergs o sentimientos.

Los rasgos dinámicos no son tan estables como los de temperamento y aptitudes. La fuerza de los sentimientos varían por el aprendizaje, la decadencia, el tiempo y las circunstancias. Los niveles de tensión érgica están cambiando debido a la estimulación general del medio y a la cantidad de reforzamiento de metas que reciben generalmente (Hall y Lindzey, 1984; Cueli y Reidl, 1972)

ESTADOS, ROLES Y DISPOSICIONES.

Los estados de ánimo son fluctuantes; la persona asume o abandona un determinado rol y se adoptan disposiciones mentales momentáneas en relación a ciertos aspectos del ambiente. Por su influencia en la conducta se incluyen en la ecuación de especificación que suma rasgos de capacidad, temperamento, ergs, estados, roles, disposiciones; cada uno con su respectivo índice situacional que define su importancia. La personalidad es aún un perfil que incluye todos los factores que pueden afectar la respuesta en un momento determinado.

Se han identificado factorialmente los estados siguientes: la ansiedad, el stress, la fatiga, la actividad autónoma general y los que corresponden posiblemente a los factores de rasgo A (sociable-reservado), F (entusiasta-sobrio) y H (audaz-tímido)

Catell también ha investigado la ansiedad considerándola un rasgo pero estos dos planteamientos rasgo-estado no son incompatibles: el nivel de ansiedad de una persona puede ser un rasgo característico y como un estado, puede fluctuar significativamente con las influencias situacionales y organizativas. (Hall y Lindzey, 1984)

LA ESFERA DE LA PERSONALIDAD ANORMAL

Se ha mencionado que el trabajo de Catell, en principio, se enfocó a la personalidad normal, pero posteriormente extendió sus investigaciones a la esfera de la personalidad anormal en su intento por abarcar la totalidad de la conducta. Debido a que la clasificación de rasgos inicial se basó en estudios de personas normales, la pregunta obligada era si estos factores correspondían a la personalidad anormal y si se habían omitido algunos factores. Si los factores fueran distintos, este hecho por sí solo haría necesario la investigación en este campo. Catell y Kline, (1982) exponen la forma en que fueron identificados los factores de la personalidad anormal. Se utilizaron tres fuentes de datos:

- Descripciones de comportamientos patológicos encontrados en textos psiquiátricos
- Nuevos elementos escritos por Catell y sus colaboradores en su práctica clínica
- La amplia gama de elementos del MMPI que fue la fuente más importante.

Se eligió el MMPI, por considerar que es útil en el estudio de anormales y que podía servir de base para identificar las dimensiones de la conducta anormal, estudiándolo a la luz del análisis factorial.

No se considera útil el factorizar las escalas clínicas por la forma en que fueron construidas; aún así, se han hecho intentos serios y los resultados muestran habitualmente que no hay más de cinco dimensiones subyacentes a las escalas clínicas del MMPI.

La alternativa de factorizar las correlaciones de los elementos es más provechosa pero el análisis factorial de una matriz de 566 correlaciones es casi imposible. Catell y Bolton optaron por agrupar los elementos en paquetes homogéneos factorialmente y procedieron a factorizar las correlaciones entre estos paquetes junto con los factores del 16 PF. Encontraron que todos los factores del 16 PF, excluyendo los siete factores nuevos y el factor B, surgían con otros cinco, que fueron interpretados como: hipocondría psicótica, comportamiento psicopático, psicastenia e interés religioso. El quinto factor no fue identificado. Este estudio en especial reveló la existencia de cuatro nuevos factores que corresponden a la conducta anormal.

En estudios posteriores se identificaron doce factores, sobre todo con datos Q, que corresponden a la esfera de la personalidad adulta anormal en la concepción de Catell y con esta base se desarrolló un nuevo instrumento, el Cuestionario de análisis clínico CAO (Catell y Krug) que mide estos doce factores patológicos además de seis factores normales. (Catell y Kline, 1982) Esta prueba está traducida al español y normalizada en España.

La concepción de Catell sobre el desarrollo de la patología se basa en un "modelo de conflicto". Este modelo explica que inicialmente se presenta la estimulación y luego el bloqueo de un impulso; la persona trata de romper esta barrera pero fracasa, de modo que se ve forzado a renunciar a ello, lo que la lleva a un estado de conflicto y ansiedad que después se convierte en un sintoma neurótico. Catell incluso llegó a calcular un índice matemático de conflicto y se comprobó que este índice permitía distinguir a los pacientes mentales de otros pacientes. Este método considera la estimación de conflicto realizada por los psiquiatras y se vincula con la debilidad del ego medida por el 16 PF (Catell, 1959 en Colom, 1995)

Para finalizar, se mencionan algunas de las propuestas fundamentales de la teoría de la personalidad de Catell revisadas por Corsini (1977 en Catell y Kline, 1982) Se han considerado las propuestas de conceptos más generales y se han dejado de lado las de carácter más técnico o más relacionadas con aspectos de motivación.

1. El análisis multivariado muestra una estructura de los rasgos fundamentales. Usa factores comunes y únicos que se encuentran en estratos distintos
2. Los rasgos poseen modalidades distintas que se corresponden con los rasgos aptitudinales, temperamentales y dinámicos (aproximadamente unos 30 en cada modalidad)
3. Las estructuras de rasgos se encuentran en distintas edades donde evidencian un desarrollo continuo. Asimismo, se han hallado y constatado en diferentes culturas.

4. Las estructuras de estos rasgos se ubican en cualquiera de los tres medios de observación: el comportamiento calificado *in situ*, los cuestionarios y las pruebas objetivas que exigen un determinado tipo de actuación.

5. La Ecuación de Especificación permite una constatación cuantitativa de la acción de los rasgos y utiliza puntuaciones factoriales ponderadas.

6. Es necesario considerar los "estados" para realizar una especificación razonable y cabal. Existen técnicas para obtener tales estados y presentan unos doce ejes tales como ansiedad, depresión, tensión-fatiga y los niveles de tensión ergica.

7. El nivel de un estado es el efecto de la acción moduladora proveniente de las situaciones ambientales sobre la proclividad de un rasgo-estado.

8. Los rasgos fundamentales están determinados en su origen, en parte por la genética y en parte por el ambiente aunque la proporción de ambos varia de rasgo a rasgo.

9. En el ambito dinámico se han identificado mediante analisis factorial, estructuras dinámicas que se han reconocido como ergios (impulsos) y sentimientos (patrones aprendidos ante instancias sociales)

10. La Ecuación de Especificación es una fórmula matemática aplicada al comportamiento que considera tanto la situación que sirve ds estímulo como la respuesta. Se trata de una constatación paramétrica múltiple en cada término en el contexto de la formulación básica de estructuras, organismo y respuesta.

11. El cambio del organismo abarca el aprendizaje, la volución (motivación e involución) y la experiencia somática.

12. El componente de aprendizaje cambia en tres vectores: en los rasgos, en el vector perceptivo-motor y en el vector involucración (emotiva) a diferencia de las teorías reflexológicas que plantean el cambio en una sola variable

Corsini se extiende en otros puntos profundizando en la teoría Catell respecto al campo dinámico o motivacional, los procesos de ajuste, el conflicto-decisión-integración, etc.

La teoría de Catell cuenta con amplios fundamentos empíricos y esto se le debe reconocer para su crédito pero para su descrédito está el hecho de que con frecuencia proclama una base empírica más estable de lo que realmente tienen sus constructos. La teoría de Catell no es precisamente popular como la de Freud, la de Rogers, Sullivan o Allport o Murray en parte por la complejidad prohibitiva del método factorial, en parte por la afirmación empírica exagerada e innecesaria y en parte por la ofensa a los enfoques alternativos. Quizá aún no sea tiempo de evaluar su trabajo y ubicarlo en su justa medida porque aún se está generando y está sufriendo una constante evolución, revisiones y cambios (Hall y Lindzey, 1984)

1.5 INVENTARIOS DE PERSONALIDAD IMPORTANTES

1.5.1 EL CUESTIONARIO 16 PF DE CATELL.

Esta prueba es el resultado de amplias investigaciones realizadas por Catell y sus colaboradores y es el más importante de sus instrumentos de evaluación por las investigaciones realizadas en torno a él. Según el investigador inglés Furham (1992) es quizá el más famoso de todos los tests de personalidad utilizados en escenarios ocupacionales, organizacionales e industriales.

La revisión precedente de la Teoría de Catell sirve de base para comprender el origen, la construcción y la estructura de la prueba. Este inventario es uno de los que cuentan con la más amplia investigación anterior y posterior a su construcción. Para plantearlo de una manera resumida, la base de la construcción de esta prueba fue una amplia lista de adjetivos descriptivos de características personales que fueron reducidas, analizadas factorialmente y sobre las dimensiones identificadas, se construyeron reactivos que a su vez fueron analizados factorialmente. El resultado es este test que evalúa 16 factores básicos y de cuatro a ocho factores secundarios.

La primera publicación comercial del 16 PF tuvo lugar en 1949 y le precedieron diez años de investigación empírica de análisis factorial. De entonces a la fecha no han cesado las investigaciones en torno a este instrumento del cual se han publicado diferentes revisiones. Su uso se ha extendido por países de América, Europa, Australia, etc.

En España, la adaptación del 16 PF fue publicada en 1975; existe una adaptación para directivos (Seisdedos, 1981) y una actualización psicométrica de la misma (Seisdedos, 1990). Se disponen de investigaciones adicionales acerca de su utilidad y aplicación clínica (Fernández-Ballesteros, 1992). Existe una nueva versión publicada en 1993 en EE.UU., el 16 PF-5 (Catell et al), traducida a nuestro idioma y adaptada en España en 1995 de la cual no se ha realizado aún ninguna adaptación en México.

Esta descripción se centrará principalmente en la versión en español publicada en 1980 por el Manual Moderno (que corresponde a la versión de 1970 de EEUU) por ser ésta la versión que es factible adquirir y la que circula actualmente en nuestro medio.

LOS CRITERIOS QUE RECOMIENDAN EL USO DEL CUESTIONARIO 16 PF (Catell, Eber y Tatsuoka, 1980; 1988):

- a) Esta inserto en un amplio tejido de la teoría psicológica general. La investigación previa de la esfera del lenguaje, la identificación de factores, la construcción de la prueba, la reconfirmación de los factores, las revisiones de la prueba, las investigaciones posteriores, etc. dan como resultado este test.
- b) En su forma actual, 1970 en Estados Unidos y 1980 en México, la prueba se basa en más de 10 investigaciones de análisis factorial sobre una muestra de varios miles de reactivos. (No se encontraron datos de análisis factoriales realizados en México)
- c) Las propiedades psicométricas de las escalas, por ej. confiabilidad y validez, han sido exploradas y publicadas para una gran variedad de muestras y condiciones fuera de México.

d) Los datos científicos acerca del cuestionario proporcionan una rica base de evidencia y de criterio en Psicología industrial, clínica, social y educacional

e) Las escalas son independientes por lo que las correlaciones entre ellas son realmente bajas.

f) Estabilidad intercultural del rasgo. Se ha comprobado la universalidad de la estructura del factor como evidencia de una estructura básica similar del rasgo original de las personalidad entre las culturas. Existen estandarizaciones del 16 PF al menos en 15 países diferentes como Francia, Italia, Alemania, Japón, Reino Unido

g) La continuidad de la estructura del rasgo. Los rasgos principales identificados por el 16 PF persisten a través de una secuencia de desarrollo. Se ha encontrado una estructura de rasgos similares entre niños pequeños, niños, adolescentes y adultos como lo demuestran las investigaciones con los tests para diferentes edades

DESCRIPCION DEL INSTRUMENTO

La prueba original en inglés cuenta con cinco formas y las diferencias entre estas están en función de los propósitos del examen. La forma A es la versión corriente para adultos y la forma B es la alternativa equivalente. Las formas C y D son más cortas y fueron diseñadas para una aplicación rápida con fines de selección, por ejemplo. Estas últimas cuentan con una Escala de Distorsión Motivacional original para detectar el intento de presentar una buena imagen lo cual es especialmente útil en situaciones de evaluación de personal. Por otro lado, la Forma E se diseñó para ser aplicada a sujetos con una capacidad inferior de lectura al esperado para el nivel medio de educación en la sociedad

EL 16 PF EN ESPAÑOL

En España se dispone de las formas A, B y C con ponderaciones obtenidas en muestras españolas. (Karson y O'Dell, 1989) En México, la prueba fue editada en 1980 por El Manual Moderno e incluyen las formas A, B, C y D pero solamente se dispone de baremos para la forma A. Estas formas no cuentan originalmente con escalas de validez. Anteriormente circularon versiones del 16PF editadas por: Psicología Industrial de Colombia, Psicología Industrial de Venezuela e Ipimex en México. Estas correspondían a la forma C.

¿QUE MIDE?

La versión del 16 PF de 1980 de Ed. El Manual Moderno, mide 16 rasgos primarios de la personalidad y cuatro secundarios. Los 16 rasgos primarios corresponden a los mencionados en el punto anterior, en el cuadro de Descripción de rasgos primarios normales. Son 14 rasgos de temperamento, uno de motivación, Q4 y uno de aptitud, la escala B. Se incluyó esta variable aptitudinal para completar el conjunto de datos con un rasgo de aptitud general que es importante en la mayoría de las predicciones y que se presenta como una dimensión fundamental en las diferencias individuales. (Seisdedos, 1995) Los cuatro rasgos secundarios que mide esta versión son: Extroversión, Ansiedad, Tenacidad e Independencia

MATERIAL Y DISEÑO DE LA PRUEBA.

La prueba es un cuestionario de 187 preguntas que se ordenaron ciclicamente siguiendo un plan que mantiene el interés del evaluando y facilita la calificación. Las preguntas de las formas A y B admiten tres opciones de respuesta: cierto, falso e intermedio. Los 16 factores puntúan en un continuo de dos polos con 10 grados de puntajes convertidos denominados estenes

La prueba consta de cuadernillos de aplicación, utilizables más de una vez, y una hoja de respuestas individual con un formato de perfil de las 16 escalas primarias en el reverso que facilita la calificación manual para la cual se dispone de plantillas. Se cuenta con un manual e instructivo así como tablas y datos psicométricos.

APLICACION Y CALIFICACION

Forma A y B. Las instrucciones impresas en el formato de aplicación son sencillas y fáciles de seguir. Aunque la prueba puede ser autoaplicable, se sugiere la presencia del examinador sobre todo para enfatizar la importancia de la sinceridad y la franqueza del examinando. El examinador usualmente leerá las instrucciones con el o los evaluandos y aclarará las dudas. Una alternativa es utilizar grabaciones con las instrucciones. Sea cual fuere el método a utilizar, se recomienda el mantener la estandarización de las condiciones de aplicación.

La prueba puede ser administrada individualmente, autoaplicarse o aplicarse grupalmente. Las condiciones de aplicación son las usuales para pruebas de lápiz y papel en cuanto a privacidad, iluminación, comodidad etc., así como el estado habitual del examinando. La escolaridad mínima es el equivalente a secundaria en nuestro país. Existen programas computarizados de aplicación y calificación.

El protocolo de respuestas cuenta con una breve ficha de identificación que el examinando debe llenar. La prueba no tiene tiempo límite pero normalmente toma entre 45 y 60 minutos. El aplicador debe apoyar para que los evaluandos no se rezaguen demasiado y sigan el procedimiento correcto al contestar: dar una sola respuesta a cada pregunta y asentarla correctamente en el número correspondiente de la hoja de respuestas. En aplicaciones grupales, el examinador debe asegurarse que todos los examinandos entendieron las instrucciones y debe evitar la comunicación entre ellos.

CALIFICACION

En forma manual, la prueba se califica con plantillas y se obtiene un puntaje bruto que se convierte a estenes en las tablas correspondientes. Con estas calificaciones se dibuja un perfil que sirve de base para la interpretación. Los puntajes de los factores secundarios se obtienen diferencialmente por sexos siguiendo un formato con instrucciones. Estos puntajes se expresan en estenes (Catell, Eber y Tatsuoka, 1980)

Se disponen de algunas versiones computarizadas que proporcionan una calificación rápida, que ahorran tiempo, evitan errores y son bastante recomendables en el caso de un gran número de aplicaciones. Si se opta por una calificación computarizada, deben seguirse las instrucciones para los protocolos de aplicación. Otros programas ofrecen además de la calificación, la interpretación de los resultados con un informe narrativo pero no se tienen suficientes datos para evaluar la precisión de estas interpretaciones.

ESTENES. Llamados decatipos por los españoles. Las puntuaciones brutas deben convertirse a estenes; esta es una calificación normalizada del 1 al 10, con una media de 5.5 y una desviación estándar de 2. Las puntuaciones directas no tienen significado por sí mismas, sólo son interpretadas en función de los estenes obtenidos de grupos normativos. Los valores 5 y 6 comprenden el 40% medio en una distribución normal y el rango de los valores 4 al 7 comprenden el 68.26 %, o sea, una desviación estándar al lado de la media. Los puntajes 2, 3 y 8, 9 son fuertes desviaciones en tanto que los valores 1 y 10 son alejamientos extremos que aparecieron muy raramente en los grupos normativos (Karson y O'Dell, 1989)

Existen dos formas de calcular los estenes; en el primer caso, los llamados *estenes-s* se calculan a partir de la puntuación en bruto pero esto tiene el riesgo de producir distribuciones asimétricas en las que la media y la mediana no coinciden y en algunas ocasiones las puntuaciones extremas, 1 y 10, no encajan en la escala de desviaciones estándar.

Para evitar esos inconvenientes se utiliza el denominado *esten-n* cuyas fronteras cortan el mismo porcentaje de la población en la curva de distribución de puntuaciones en bruto dada que la que trazarían intervalos sucesivos de media sigma en una curva perfectamente normal. Esto garantiza que habrá relaciones exactas y constantes entre los puntos que cortan los estenes y las puntuaciones en centiles; entonces es fácil convertir la escala B expresada en *esten-n* a centil y centil a Cociente Intelectual, si se requiriera. (Catell, Eber y Tatsuoka, 1980)

HOMOLOGACION DE PUNTAJES

El manual del 16 PF (1980), presenta una figura de distribución normal que permite convertir diferentes puntajes, de *esten-n* a centil, a C.I. clásico o de desviación y que es especialmente útil en el caso de que el 16 PF forme parte de una batería de pruebas y sea necesario homologar los puntajes. (p 101)

La conversión de puntajes en bruto al sistema de estenes implica valores decimales y por conveniencia se usan valores enteros redondeados pero si se requiere, es factible calcular el valor exacto del *esten* por la interpolación en el caso de los *estenes-n* o a partir de la media y la sigma de las puntuaciones en bruto que proporciona el suplemento del manual, en el caso del *estenes-s*.

NORMAS Un aspecto de primordial importancia es el elegir o contar con las normas de conversión adecuadas. Se ha determinado, mediante múltiples investigaciones, la necesidad de contar con normas diferenciales por sexo (Catell, Eber y Tatsuoka, 1980) Otra necesidad indispensable es el contar con normas para el país y en el idioma en que se está utilizando el instrumento así como con una actualización en el tiempo para mantener la vigencia de la comparación.

Desafortunadamente, este instrumento no cuenta con una normalización reciente en México y se carece por completo de baremos para las forma B, C y D. El manual presenta normas para adultos mexicanos hombres y mujeres, pero al parecer en muestras reducidas. En el caso de las mujeres es de $n = 49$ y en el caso de los hombres no se registra el número. El manual tampoco refiere la fecha ni el lugar en que fueron obtenidos tales baremos referidos por Krug en 1971. (Catell, Eber y Tatsuoka, 1980)

INTERPRETACION

El manual ofrece varios datos para cada factor: la letra, el número y el nombre del factor, por ej. 1. Factor A, EXPRESIVIDAD EMOCIONAL, el término técnico, *sizothimia vs Affectothymia* y los datos que originaron esta escala, (L+ Q). Se proporciona además un conjunto de adjetivos de uso cotidiano que describen el factor y que son útiles para describirlo con las personas en general. Los adjetivos también se ubican en los dos extremos de la escala bipolar.

-El primer paso en la interpretación consiste en identificar las escalas con estenes altos (8-10) o bajos (1-3) y expresarlas con los términos que identifican a los factores en cada extremo del polo.

-Los resultados también pueden interpretarse por medio de la técnica de perfiles que consiste en comparar el perfil del sujeto con los perfiles obtenidos de grupos seleccionados. (Catell, Eber y Tatsuoka, 1980) Esta aplicación se extiende a casos de perfiles "ideales" como en el campo de la selección de personal en el que se identifica el perfil del empleado exitoso y se compara con los candidatos a tal puesto. Al comparar visualmente los dos perfiles, se localizan concordancias y diferencias para interpretarlas según el propósito de la evaluación.

-Adicionalmente a la comparación de perfiles, puede realizarse una interpretación "psicológica" intuitiva a partir del conocimiento de los rasgos fundamentales del cliente y la naturaleza de la situación. Esto daría como resultado un juicio parcialmente intuitivo más que un planteamiento alcanzado cuantitativamente. El psicólogo experimentado puede enriquecer la apreciación estadística con los juicios basados en el conocimiento psicológico de los factores y la comprensión de las leyes que operan en ciertas circunstancias dadas. (Catell, Eber y Tatsuoka, 1988)

-Otra técnica interpretativa conocida como cálculo de criterios, más compleja, comprende el uso de ecuaciones estadísticas en las que cada calificación se multiplica por el valor relativo numérico específico a fin de predecir ciertos comportamientos o clasificaciones. (Catell, Eber y Tatsuoka, 1980)

-Otro procedimiento de interpretación se inicia con el análisis de las calificaciones obtenidas en las escalas de validez para determinar si las calificaciones son lo suficientemente válidas para garantizar la interpretación y, en ese caso, se procede a analizar las calificaciones de los factores de segundo orden seguidas por las de los 16 factores. (Karson y O'Dell en Aiken, 1996) En mi práctica profesional me ha sido útil la Escala B como una alternativa adicional de validez: si el puntaje es demasiado bajo, debe considerarse si el grado de comprensión de información escrita del evaluado garantiza las respuestas proporcionadas a los reactivos.

-Se ha utilizado el 16 PF para identificar Tipos de temperamento dentro de un marco multidimensional o de rasgos múltiples. Boltz (1976 en Buss y Poley, 1979) ha identificado seis tipos obtenidos a partir de cuidadosas investigaciones con el 16 PF. Aunque este enfoque ha sido relativamente poco estudiado, ofrece una opción interesante de interpretación.

-El manual indica otros métodos de interpretación cuantitativa que requiere desarrollar algunos cálculos y seguir formulas más o menos complicadas. Esto permite una comparación muy precisa entre los perfiles de individuo a individuo, de individuo a grupo y de grupo a grupo. Estas comparaciones pueden hacerse considerando todos los factores o solamente algunos, según el propósito que se persiga. (Catell, Eber y Tatsuoka, 1980) Esta técnica es especialmente útil en la evaluación de grupos numerosos si se cuenta con apoyo de computadora y se utiliza más bien en labores de investigación en donde la precisión de las diferencias es de primordial importancia.

VALIDEZ Y CONFIABILIDAD

Validez. (Catell, Eber y Tatsuoka, 1980) Los reactivos del cuestionario fueron elegidos por su validez significativa comparados con los factores después de diez análisis factoriales sucesivos y fueron seleccionados entre varios miles de reactivos probados originalmente. Asimismo, se considera que la validez de la prueba misma es un concepto o "construcción" de validez. Se incrementa la validez, lo mismo que la confiabilidad, usando más de una forma de la prueba.

Los resultados de los cálculos del concepto directo de validez (la correlación entre la puntuación de la escala con el factor puro que estaba diseñado para medir) arrojaron los resultados siguientes:

Para las formas A + B = de .53 a .94 entre las diferentes escalas.
Para las formas C + D = de .45 a .91 "
Para la forma A = de .35 a .92 "
Para la forma B = de .44 a .87 "

CONFIABILIDAD

Según su autor, varía en las formas A y B de 0.45 a 0.93 y en las formas C y D, de 0.67 a 0.86, pero esto ha sido discutido por otros autores. La adaptación española de la prueba alcanza unas correlaciones de reaplicación en periodos de 1 a 4 semanas de 0.62 a 0.78 según diferentes factores. Existen asimismo numerosas críticas sobre la fiabilidad de este instrumento (Catell y Johnson, 1986, Seisdedos, 1990, Furham, 1990 en Fernández-Ballesteros, 1992) Por supuesto, la confiabilidad aumenta si se aplica más de una forma de la prueba.

Los coeficientes de confiabilidad de estabilidad temporal, reaplicación en intervalos cortos, referidos por el Manual (Catell, Eber y Tatsuoka, 1980), varían en la forma A = de .72 a .92 en las diferentes escalas y en la forma B = de .54 a .89 en las diferentes escalas.

La reapiación en intervalos de 2 a 24 meses, dió como resultado los siguientes coeficientes de estabilidad temporal:

Formas A + B con A = de .63 a .87 en las diferentes escalas.

Forma A con B = de .35 a .85 en las diferentes escalas

Los coeficientes de equivalencia, formas A con B = de .21 a .71 en las diferentes escalas.

formas A con C = de .16 a .55 en las diferentes escalas.

Notése que en algunas escalas los coeficientes son bastante bajos.

La confiabilidad baja a moderada del 16 PF es una desventaja importante. La confiabilidad por división por mitades de .54 y los coeficientes de equivalencia (las correlaciones entre las escalas de factores primarios de las diferentes versiones de la prueba), también son bajos. Los coeficientes de test-retest durante los intervalos breves de tiempo (coeficientes responsables) se encuentran entre .70 y .80 lo cual significa que son más bajos que los de otros inventarios de personalidad. (Aiken, 1996)

En España la confiabilidad o coeficiente de equivalencia entre la forma A y la forma B, arrojó resultados de .14 a .78 y el coeficiente test-retest arroja resultados de .52 a .78 (Seisdedos, 1985)

Los datos del 16 PF aplicado en escenarios laborales no son concluyentes. Furham (1991 en Furham, 1992) no encontró capacidad predictiva del 16 PF para discriminar entre un grupo de tripulación de cabina de una línea aérea clasificado por otros criterios en tres categorías: empleados buenos, regulares y malos. Bernardin (1977 en Furham, 1992) por su parte, encontró correlaciones moderadas entre dos escalas del 16 PF y el ausentismo y la separación del empleo. Moore y Stewart (1989 en Furham, 1992) encontraron varias correlaciones entre el 16 PF y la honestidad e integridad de empleados.

1.5.2 EL INVENTARIO MULTIFASICO DE LA PERSONALIDAD DE MINNESOTA, MMPI.

Esta prueba construida por Hathaway y McKinley, se publicó originalmente en 1943. Algunos autores lo consideran el cuestionario de personalidad que se ha estudiado más a fondo y que mejor ejemplifica el enfoque psicométrico. Se ha convertido en la base para la investigación de muchos rasgos y tipos de personalidad, se ha utilizado extensamente y se ha escrito ampliamente acerca de él. Se le considera la prueba de personalidad más utilizada tanto en ambientes clínicos como normales.

Contiene varias escalas de autoinforme que originalmente se diseñaron para clasificar a los pacientes con trastornos mentales.

CONSTRUCCION POR CRITERIO EMPIRICO.

Este inventario no fue construido con analisis factorial sino con la técnica denominada de Criterio empirico. Estos tests se forman con reactivos que diferencian entre dos o más grupos de criterio. Los autores del MMPI seleccionaron aquellos items que resultaron con mayor poder de discriminación entre diferentes grupos psiquiátricos y normales. (Fernández-Ballesteros, 1992)

DESCRIPCION Y ESCALAS.

En este punto se presentará una breve descripción del instrumento (Aiken, 1996, Fernández-Ballesteros, 1992 y Nuñez, 1990), sin profundizar en su construcción ni en su interpretación que es muy amplia. Esta descripción servirá de base para analizar sus escalas de validez en el siguiente capítulo.

El MMPI original se diseñó para identificar alguna anomalía psicológica. Las afirmaciones están relacionadas con actitudes, emociones, incapacidades motrices, síntomas psicósomáticos además de otros sentimientos y comportamientos que indican problemas psiquiátricos. Las escalas se formaron con reactivos que distinguan entre las respuestas de un grupo de pacientes psiquiátricos específico y un grupo control de personas normales. Estas escalas se validaron al comparar las respuestas de los pacientes psiquiátricos con aquellas de distintos grupos de individuos normales. (Aiken, 1996) Específicamente Catell tomó el MMPI como base para obtener información en sus estudios en la personalidad anormal, como se mencionó anteriormente.

En su forma final, contiene 566 items de respuesta falso-verdadero. Los resultados se expresan en puntajes T, con media de 50 y desviación estándar de 10. Esta prueba está ampliamente estandarizada en muchos países. Una versión corta utiliza 399 reactivos que se refieren a las 10 escalas básicas y a las escalas de validez.

Escalas. La primera forma del MMPI constaba de ocho escalas pero posteriormente se agregaron cuatro escalas de validez para evaluar el grado y tipo de distorsión en el autoinforme y otras dos escalas, la de Masculinidad-feminidad y la de Introversión social. También había escalas adicionales. Un gran número de escalas se fueron desarrollando paralelamente al extenderse el uso del instrumento.

Las escalas clínicas quedaron de la forma siguiente a partir de 1948 y es el formato que se ha utilizado a través de los años y que continúa vigente en la revisión del instrumento

1. Hipocondrías	6. Paranoia
2. Depresión	7. Psicastenia
3. Histeria	8. Esquizofrenia
4. Desviación psicopática	9. Manía
5. Masculinidad-feminidad	0. Introversión-extroversión

ESCALAS DE VALIDEZ DEL MMPI

Esta prueba cuenta con cuatro indicadores de la disposición o actitud del sujeto al contestar a los reactivos. (Núñez, 1994) Los puntajes obtenidos y las distintas combinaciones entre ellos proporcionan información valiosa para la interpretación de los resultados en las escalas clínicas

Escala de frases omitidas (?) es el número de preguntas que el sujeto no contestó.

Escala de mentira, L. Indica el grado de mentira o franqueza del sujeto.

Escala F. De validez, evalúa las respuestas infrecuentes

Escala K. De defensa. Una fracción del valor de la escala K se utiliza como corrección en las escalas clínicas 1, 4, 7, 8 y 9.

El tema de las escalas de validez del MMPI se ampliará en el capítulo 3.

INTERPRETACION.

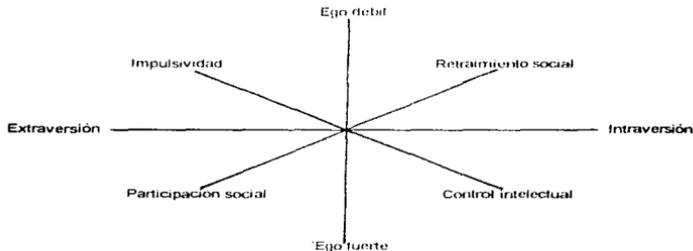
Como primer paso es indispensable evaluar las escalas de validez de la prueba para determinar precisamente el valor de la información que el examinado presenta de sí mismo.

La relación entre las puntuaciones más altas obtenidas por cada sujeto en las diferentes escalas básicas resulta en perfiles que se interpretan en el juicio diagnóstico. Los puntajes T superiores a 70 se consideran indicadores de patología, en primera instancia, pero la interpretación es bastante compleja y amplia. Los códigos formados por dos o tres de las escalas con mayor puntuación han sido exhaustivamente estudiados. Desde 1965 existen *software* para la interpretación computarizada. (Fernández-Ballesteros, 1992) La obra de Rivera (1991) proporciona amplia información de la interpretación del MMPI y uno de sus principales aciertos, en mi opinión, es que el lenguaje empleado es comprensible y aplicable a informes de evaluación de personal.

ANALISIS FACTORIALES DEL MMPI

Se han hecho análisis factoriales de sus escalas clínicas y se han obtenido dos grandes factores de primer orden: Ansiedad y Represión (Welsh, 1956 en Fernández-Ballesteros, 1992). Otros autores como Kassenbaum et al. 1959, y Korah, 1964 (citados por Eysenck, 1969) realizaron análisis factoriales del mismo instrumento y obtuvieron cuatro factores que se muestran en la siguiente figura pero ni éstos ni los resultados de otros análisis factoriales fueron tomados en cuenta para la reestructuración de la prueba a pesar de su relativa coincidencia con los factores de segundo orden de Catell, con los de Norman y los de Eysenck. (en Fernández-Ballesteros, 1992)

DIMENSIONES DE LA PERSONALIDAD POR ANALISIS FACTORIAL DEL MMPI
Identificadas por Kassenbaum et al. 1959, Korah, 1964, en Fernández-Ballesteros, 1992



Costa y McCrae (1980 en Morris, 1992) también analizaron factorialmente al MMPI en una amplia población y no encontraron factores significativos que difieran de las características normales de la personalidad.

Algunas de las críticas más duras en contra del MMPI provienen de Eysenck (1991 en Fernández-Ballesteros, 1992): los resultados del análisis factorial de las escalas hipotetizadas indican que los ítems correlacionan más con escalas a las que no pertenecen que con las que se integran. Considera que el inventario de personalidad más usado viola todas las reglas dadas por los psicometras para la construcción de este tipo de instrumentos y que esto podría deberse a que se basa en una oscura teoría de personalidad, que las escalas resultantes pueden ser interpretadas con alta subjetividad y basadas en categorías carentes de significación científica. Además, estas categorías son diagnosticadas con una extrema falta de fiabilidad.

EL MMPI-2.

En 1989 se publicó una revisión del **MMPI** (Bucher et al, 1989 en Fernández-Ballesteros, 1992), que perseguía los siguientes objetivos: mantener las escalas originales, actualizar las normas (lo "normal" ya no es lo mismo), incluir contenidos que no se representaban en la versión original y revisar la redacción de las preguntas cuyo lenguaje estaba pasado de moda, era inconveniente o mostraba sesgo para algún sexo. Esta revisión, en idioma inglés, ofrece versiones diferentes para adultos y adolescentes.

La versión para adultos contiene 154 reactivos experimentales nuevos, para evaluar áreas de la psicopatología que no estaban bien representadas en el **MMPI** original, como trastornos alimenticios, personalidad Tipo A y abuso de drogas.

La versión para adolescentes contiene 104 reactivos nuevos relacionados con problemas específicos de la adolescencia.

Esta nueva versión se diseñó para propósitos clínicos y no clínicos pero se está utilizando más para evaluar la personalidad normal.

El **MMPI-2** contiene 567 preguntas de falso y verdadero. Las cuatro escalas de validez y las diez escalas clínicas básicas se califican con los primeros 370 reactivos y el resto proporciona material para las escalas de contenido e investigación muchas de las cuales se conservan del **MMPI** original, en total 18 escalas complementarias, o escalas de contenido y 38 subescalas.

Otro cambio importante es en relación a los puntajes *T* que se hicieron uniformes en ocho de las escalas clínicas y otras escalas de contenido. Las calificaciones *T* tradicionales no eran estrictamente comparables. Las calificaciones *T* uniformes son comparables y conservan la forma general de las distribuciones originales de las calificaciones *T*.

Esta nueva versión se estandarizó en una muestra de 2,600 residentes estadounidenses de 18 a 90 años de edad seleccionada con criterios del censo de 1980 para garantizar la representatividad. Los datos de confiabilidad se derivan de grupos relativamente pequeños: $n=193$; y arrojan coeficientes de reaplicación de las escalas básicas que varían de .58 a .92. Deben tomarse con cautela las diferencias en las calificaciones obtenidas en las escalas debido a los bajos coeficientes de confiabilidad aunados a los considerables errores estándar de medición. En esta nueva versión las escalas se identifican con el número que tienen asignado.

En ambas versiones, la interpretación se inicia con una codificación de las escalas clínicas y se presta especial atención a las escalas con calificaciones *T* más altas. (Aiken, 1996; Fernández-Ballesteros, 1992)

Aunque el **MMPI** original es el inventario de personalidad más ampliamente usado, las escalas carecen de significación psicológica por haber sido construidas con un criterio empírico, lo mismo puede aplicarse al **MMPI-2**. Lo más que puede decirse de tales escalas es que discriminan o no entre grupos-criterio particulares. (Kline, 1993b)

CAPITULO 2.

PRINCIPALES ENFOQUES PARA LA EVALUACION DE LA PERSONALIDAD

En este capítulo, se revisará el concepto de Evaluación Psicológica su ubicación en la Psicología, los modelos de Evaluación Psicológica y se describirán los dos principales enfoques para la evaluación de la personalidad: el psicométrico y el proyectivo. Se finalizará con un análisis de los inventarios de personalidad como instrumentos representativos del enfoque psicométrico para la evaluación de la personalidad.

En el capítulo anterior se refirieron las definiciones de personalidad de Eysenck, de Guilford y de Catell y se precisó que en las tres se consideraba, de una u otra manera, la totalidad de la conducta o la suma total de los rasgos. En este trabajo se va a considerar fuera de esta definición a los rasgos de capacidad dado que forman un grupo de rasgos definido y distintivo. Se considerará la personalidad como un conjunto de rasgos de temperamento y dinámicos, con referencia al modelo psicométrico (Kline, 1985). Tradicionalmente el término de pruebas de capacidad ha sido utilizado para instrumentos que evalúan habilidades o rendimientos en las áreas cognoscitiva, intelectual y motora. Cronbach sugirió que se considerara a las pruebas de capacidad como pruebas de ejecución máxima y a las pruebas de personalidad como pruebas de ejecución habitual o instrumentos que se espera que identifiquen los modos típicos de responder del individuo (Lanyon y Goodstein, 1977). Una definición diferente de personalidad se incluíra dentro del análisis de las pruebas proyectivas.

Al tratar de conjugar una explicación sobre la personalidad y un método para evaluarla, el problema se torna doblemente complejo. "la interacción entre la conceptualización de un constructo y su evaluación es un proceso reverberante" (Fernández-Ballesteros, 1992)

2.1 LA EVALUACION PSICOLOGICA

La Evaluación psicológica es una subdisciplina de la Psicología científica. La base teórica y metodológica de la Evaluación psicológica la proporciona la Psicología diferencial y la Psicología de la personalidad pero también se vincula y ha recibido aportaciones importantes de la Psicopatología, la Psiconeurología, la Psicología social y las aplicaciones de la Psicología del aprendizaje a la Clínica. La metodología desarrollada por la Psicología experimental le ha sido de gran utilidad aunque también ha ido construyendo su propia y específica metodología.

Las exigencias sociales y las necesidades en el ámbito clínico, escolar y de las organizaciones han dado lugar al desarrollo de la tecnología de la evaluación psicológica y le han conferido la categoría de una disciplina eminentemente aplicada. (Fernández-Ballesteros, 1992)

Una diferencia fundamental entre la Psicología y la Evaluación psicológica estriba en la generalidad de su aplicación. La Psicología estudia la conducta humana para encontrar los principios generales que en ella rigen mientras que la evaluación se dirige a un sujeto, un ente individual o, por extrapolación, a un grupo.

El objetivo de la evaluación es verificar si los principios generales de la Psicología en sus distintas especialidades se dan en ese sujeto particular. Se define como "aquella disciplina de la Psicología que se ocupa del estudio científico del comportamiento (a los niveles de complejidad necesarios) de un sujeto (o un grupo) con el fin de describir, clasificar, predecir y, en su caso, explicar y controlar tal conducta" (Fernández-Ballesteros, 1992) Lanyon y Goodstein (1977) la describen como el proceso de recoger y organizar la información acerca de un individuo con la expectativa de que esta información conducirá a una mejor comprensión de esa persona lo que implica el hacer alguna predicción sobre su conducta en términos específicos.

La Evaluación Psicológica cuenta con una metodología particular y se orienta a objetivos aplicados de descripción, diagnóstico, orientación, tratamiento, decisiones organizacionales, etc. Como en la ciencia básica, existen diferentes enfoques o modelos.

2.1.1 MODELOS DE EVALUACION PSICOLOGICA.

La preocupación por cuestionarse el objeto y el proceso de la evaluación a partir de los setentas lleva a definir una serie de modelos que Fernández Ballesteros (1992) presenta, vinculándolos con una investigación realizada en 1968 por Coan quien se propuso investigar acerca de las tendencias constantes a lo largo del tiempo y de los sistemas psicológicos con el fin de identificar las dimensiones más importantes presentes en la Psicología. El análisis las variables propuestas por 54 importantes teóricos y detectó 34 variables de contenido, metodología, supuestos básicos y modos de conceptualización, las que sometió a análisis factorial. Sus hallazgos pusieron de manifiesto seis factores bipolares: endógeno-exógeno, dinámico-estático, mentalista-objetivo, molar-molecular, nomotético-ideográfico y cuantitativo-cualitativo.

Esos seis factores están presentes -entre otros que se consideraron igualmente importantes- en los modelos considerados que son los siguientes:

- | | |
|-----------------------------------|-------------------------|
| 1. El modelo del Rasgo o Atributo | 2. El modelo dinámico |
| 3. El modelo médico | 4. El modelo conductual |
| 5. El modelo cognitivo. | |

El análisis de estos modelos se basa en varios criterios: su formulación teórica y los supuestos sobre los determinantes de la conductas -en referencia al factor endógeno-exógeno de Coan-, esto es, si la conducta está en función de variables internas o externas. También considera el nivel de inferencia, la clase de variables objeto de análisis en cada uno de los modelos, los métodos utilizados, las técnicas de evaluación características de cada uno, los objetivos que se persiguen y los ámbitos de aplicación. Aunque Fernández-Ballesteros (1992) analiza cuidadosamente cada uno de estos modelos, aquí se referirán sus características de un modo resumido y según lo relevante a este trabajo.

MODELOS EN EVALUACIÓN PSICOLÓGICA					
MODELO	VARIABLES	MÉTODOS	TECNICAS ^{ej}	OBJETIVOS	AMBITOS DE APLICACION
DEL RASGO (o Atributo)	rasgos factores dimensionales	Estrategias correlaciona- les	Cuestion de personalidad Tests de inteligencia	Predicción Descripción Señal/explica- ciones	Escolar Organizacional
DINAMICO	Construcciones teóricas Estructura de la personalidad	Observación clínica	Técnicas proyectivas	Explicación Comprensión	Clinico
MEDICO	Entidades no- sológicas Trastornos bio- lógicos	Estrategias correlaciona- les	Tests Cuestionarios Escala de observación Baterías neu- ropsicológicas	Descripción Clasificación Pronóstico Explicación	Clinico
CONDUCTUAL (radical)	Ambientales Motoras	Estrategias experimenta- les y correla- ciones	Técnicas de observación Autoinformes Registros fi- siológicos	Explicación funcional Control	Clinico Escolar Organizacional Laboratorio
CONDUCTUAL (cognitivo- conductual)	Cognitivas Motoras Fisiológicas Ambientales	Estrategias experimenta- les y correla- cionales	Técnicas de observación Autoinformes Registros fi- siológicos	Explicación funcional Control	Clinico Escolar Organizacional Laboratorio
COGNITIVO	Representa- ciones Fases de procesamiento Estrategias	Estrategias experimenta- les y correla- cionales	Ejecución ante tareas cognitivas Autoinformes	Descripción Predicción Explicación	Laboratorio Escolar Clinico

Adaptado de Fernández-Ballesteros (1992) por M. Fuentes

MODELOS DE EVALUACION PSICOLOGICA

1. El Modelo del Rasgo (o atributo). Supone que la conducta está en función de variables personales también llamadas genotípicas u organizmicas. Se usan las técnicas correlacionales para asociar tales variables internas y sus manifestaciones externas. Aquí se analizan variables intrapsíquicas obtenidas mediante procedimientos empíricos, factoriales o racionales utilizando un análisis molecular. La evaluación tiene por objeto identificar características o atributos (rasgos, factores) en que las personas difieren y se encamina más bien a la predicción y la descripción. El nombre de Atributo se lo dio Mc Reynolds pero se conoce también como el modelo psicométrico, factorial o de rasgos.
2. El Modelo Dinámico. Considera que la conducta puede ser explicada por una serie de construcciones teóricas internas que conforman la estructura de la personalidad que, aunados a una serie de dinamismos internos inconscientes, determinan la conducta. Se estudia la estructura de la personalidad desde una perspectiva molar refiriéndola a los distintos conceptos psicodinámicos. Utiliza las técnicas proyectivas y el ámbito de aplicación es principalmente clínico aunque también se utiliza en el ámbito organizacional.
3. El Modelo Médico. Explica la conducta anormal en base a factores endógenos o internos, biológicos o intrapsíquicos. Implica el estudio de cada sujeto que presenta trastornos de conducta y la forma de clasificarla según una serie de entidades nosológicas. Este modelo no se ha derivado propiamente de sistemas psicológicos sino que es una extrapolación de otras disciplinas en conjunción con las aplicaciones de la Psicología. Cuenta con sus propias técnicas de evaluación. Se utilizan estrategias correlacionales y el ámbito de aplicación es el clínico.
4. El Modelo Conductual. Plantea que el comportamiento se explica en función de variables ambientales. Desde el enfoque radical, el skinneriano, la conducta -exclusivamente motora- es funcionalmente explicada con base en sus antecedentes y consecuentes externos. Se emprende el estudio científico del sujeto mediante un análisis molecular y mecanicista de las respuestas manifiestas y de los estímulos que las mantienen. En el replanteamiento del modelo conductual-cognitivo, se incorporan las variables del organismo -o la persona- pero aceptando variables encubiertas (cognitivas) en la explicación de la conducta y considerando el ambiente en su dimensión perceptiva. Se realiza un análisis funcional entre variables ambientales y la conducta mediante técnicas observacionales, de autoinforme, etc.
5. El Modelo Cognitivo considera que la conducta puede ser explicada a través de una serie de estructuras internas, mentales, como la representación, los almacenes de memoria; y enfatiza en los procesos o estrategias cognitivas que median entre los estímulos y las respuestas o ejecuciones. Utiliza estrategias correlacionales y experimentales y se aplica en laboratorios, en ámbitos escolares y clínicos. (Fernández-Ballesteros, 1992)

Todos estos modelos permiten visualizar a la Evaluación Psicológica desde una perspectiva global. En adelante, la información se irá centrando en el tema de este estudio que es la evaluación de la personalidad en adultos, enfatizando en su aplicación laboral u organizacional.

La Evaluación Psicológica se lleva a cabo por las demandas del sujeto o de su referente y ellos son los que marcan los objetivos de la evaluación y el modelo idóneo para cubrir el propósito. En el caso de la evaluación laboral, el punto de partida es una necesidad organizacional, un puesto vacante por ejemplo, y de allí se elige el modelo, se establece el objetivo y se desarrolla el plan de evaluación.

Aunque el Modelo del Atributo se ha desarrollado ante demandas educativas y laborales y el Modelo Dinámico se ha aplicado en situaciones clínicas, ambos se han utilizado en el ámbito laboral para fines de evaluación del personal con distintos propósitos.

Coan. (1968 en Fernández-Ballesteros, 1992), en el análisis de la teoría psicológica, encontró formalmente una dicotomía que se puede denominar cualitativa-cuantitativa que está presente en los modelos presentados anteriormente. En la evaluación psicológica, el enfoque cualitativo, también llamado clínico, intenta realizar un análisis global e impresionista del sujeto. El enfoque cualitativo, como lo caracteriza Kelly (1967 en Fernández-Ballesteros, 1992), pone énfasis sobre todo en la libertad del juicio profesional. Se propugna libertad en la elección de las situaciones de examen, en la obtención y el procesamiento de los datos, en la asignación de pesos a los distintos indicadores y en la comunicación de los resultados de la evaluación. Mientras que desde el enfoque cuantitativo, también llamado psicométrico, se busca la medición de las respuestas del sujeto ante situaciones estandarizadas. Se plantea una serie de normas en cuanto a la uniformidad de los instrumentos, ponderación y procesamiento mecánico de los datos así como procedimientos uniformes de comunicación de los resultados. El margen de libertad del juicio del evaluador está bastante más restringido.

Los modelos de evaluación, especialmente el del Atributo (rasgo) y el Dinámico, sirven de base para presentar dos enfoques importantes en la evaluación de la personalidad: el Proyectivo y el Psicométrico. Estos modelos han permitido ubicar claramente las variables que manejan, los métodos, las técnicas de evaluación, los objetivos que se persiguen y el ámbito de aplicación.

En este punto, es conveniente enfatizar otra aproximación en la evaluación, implícita en los modelos: la nomotética-ideográfica; ésta se referirá aquí aplicada especialmente a los tests o técnicas de evaluación.

Los tests nomotéticos tienen que ver con el desarrollo de reglas para entender la personalidad; se dirigen a la medida de rasgos o dimensiones de la personalidad. Los cuestionarios de personalidad, tipo inventario, son medidas típicamente nomotéticas. Los tests ideográficos se dirigen a medir aquello que es único e individual de una persona, todo aquello que la persona no comparte con los demás y es esto lo que los tests proyectivos ayudan a medir. Los psicólogos clínicos, quienes trabajan con individuos, no con grupos, siempre han preferido los tests proyectivos a pesar las objeciones de los psicólogos científicos o académicos. (Kline, 1993)

2.2 EL ENFOQUE PROYECTIVO EN LA EVALUACION DE LA PERSONALIDAD

ALGUNAS CONSIDERACIONES HISTORICAS.

El desarrollo histórico de las técnicas proyectivas no puede separarse del de la Psicología en general y del de la Psicopatología en particular, entre otras disciplinas. Tampoco puede separarse del progreso de los tests en general.

Durante más de cuarenta años, las técnicas proyectivas han sido las más importantes representantes de la evaluación psicológica bien sea por la "magia" implícita en sus hallazgos que les ganó una amplia popularidad, bien sea por la evidente utilidad en el ámbito clínico.

Históricamente hay una fuerte vinculación entre la evaluación de la personalidad y la Psicopatología. Del psicoanálisis y de la Psicología dinámica se desarrollaron las técnicas proyectivas que en una época fueron las más importantes técnicas evaluativas. La ciencia avanza en un intento de resolver problemas y en el caso de la personalidad, sobre todo en los comienzos, los problemas han nacido de las psicopatologías; por esto, la evaluación de la personalidad se ha iniciado con el diagnóstico psiquiátrico pero esa evaluación ha dado lugar al surgimiento de teorías de la personalidad independientes. (Fernández-Ballesteros, 1992)

Tres serían los acontecimientos históricos más importantes en el desarrollo de las técnicas proyectivas: en 1906, Jung y su técnica de Asociación de Palabras; en 1921, la 1ª edición del psicodiagnóstico de Roschach y en 1938, la publicación de Murray de *Explorations in personality*, la base del T.A.T. (Fernández-Ballesteros, 1992) Es sobre todo a partir de los años cuarenta en que se incrementa la investigación sobre las pruebas de personalidad y, a manera de ejemplo, de los casi 5,000 estudios sobre el Rorschach hasta 1978, casi el 60% se generaron entre 1940 y 1965. En la época en que el Psicoanálisis ejerció mayor influencia sobre la teoría e investigación de la personalidad, de 1940 a 1960, fue que existió el mayor interés en el uso de las técnicas proyectivas. (Aiken, 1996)

Si ha de juzgarse por la investigación, prevalece el interés por usar este tipo de pruebas y a partir de los años ochenta se ha notado una clara recuperación de estas técnicas.

DEFINICIONES Y CARACTERISTICAS DE LAS TECNICAS PROYECTIVAS

A los instrumentos proyectivos se les conoce con la denominación de Frank de técnicas proyectivas. (Aiken, 1996) Dentro de éstas se engloban una diversidad de instrumentos de evaluación de la personalidad cuya definición no aclara si, conceptualmente, se ubican dentro de una misma clase. Se definen como los estímulos ambiguos que provocan que el sujeto pueda "proyectar" sus necesidades y sentimientos internos en respuesta a ellos bajo el supuesto de que la respuesta es un reflejo o "proyección" de la percepción individual que el sujeto tiene del mundo. (Aiken, 1996) Kline (1993) se concentra en el aspecto de la ambigüedad como lo característico de una técnica proyectiva y que es ésta lo que provoca la proyección de parte del respondiente.

Estas técnicas no suponen un grupo homogéneo de instrumentos ni por el tipo de estímulos de las pruebas, la forma de aplicar, calificar, interpretar, etc. como se verá posteriormente al describir las clasificaciones y las descripciones de las pruebas proyectivas más conocidas. Incluso en algunas de ellas la forma de aplicación y el orden de los estímulos a utilizar varían a juicio del aplicador y quizá sea precisamente esta libertad del juicio del examinador (Kelly, 1967 en Fernández-Ballesteros, 1992) presente en todo el proceso de evaluación aunada a la libertad en la respuesta del sujeto al estímulo, las características que definitivamente los engloben en una categoría definida. Sobre esta base, me permito proponer una definición alternativa de las técnicas proyectivas:

Las técnicas proyectivas son instrumentos de evaluación de la personalidad caracterizados por la libertad: libertad en la elección o conformación del estímulo, libertad en la respuesta a éste de parte del sujeto y libertad al psicólogo especialista en la interpretación de las respuestas, a la luz de un marco teórico definido.

Lindzey (1961 en Fernández-Ballesteros) define a las técnicas proyectivas como "los instrumentos considerados especialmente sensibles para revelar aspectos inconscientes de la persona que provocan una amplia variedad de respuestas subjetivas; son altamente multidimensionales y evocadores de datos inusualmente ricos con un mínimo de conocimiento por parte del sujeto del objetivo que el test persigue. Generalmente, el material de estímulo presentado es ambiguo, promueve la imaginación del sujeto, de ellos no se derivan respuestas correctas o incorrectas y su interpretación depende de un análisis global". Esta es una definición más amplia y descriptiva que la primera presentada pero persiste la duda en relación a englobar a todas las técnicas proyectivas dentro de una misma categoría.

Kline (1985) indica que las técnicas proyectivas pretenden realizar una evaluación "profunda" de la personalidad y revelar incluso aspectos de los cuales el individuo no esté "consciente"; tratan de obtener un proceso inconsciente por lo que la interpretación de las respuestas al material de estas pruebas se basa mayormente en la teoría psicoanalítica y otras teorías dinámicas. Los tests proyectivos investigan todos aquellos aspectos de una persona que son particulares de ella, la esencia de su propia individualidad.

Los psicólogos que utilizan las pruebas proyectivas intentan formar una impresión general de la personalidad del sujeto al buscar consistencias y rasgos sobresalientes en el patrón de respuestas. Como consecuencia, el uso de una prueba proyectiva típica requiere de más capacidad, sensibilidad y entrenamiento que un inventario de personalidad. (Aiken, 1996)

Se han desarrollado diferentes técnicas proyectivas a fin de detectar motivos menos obvios, conflictos, problemas y otras características intrapersonales ocultas. Las técnicas proyectivas en general son menos obvias en su objetivo y por lo tanto, se supone que están menos sujetas a la falsificación y grupos de respuestas. (Aiken, 1996; Kline, 1985) pero esta resistencia al falseamiento no es concluyente, como se explicará después.

LA BASE TEORICA DE LAS TECNICAS PROYECTIVAS.

Una de las objeciones de Eysenck (1959 en Kline, 1993) hacia los tests proyectivos es que éstos carecen de una teoría coherente. Sostiene que no existe una explicación teórica de cómo, por qué o bajo qué condiciones un individuo proyectaría algo acerca de sí mismo a partir de un estímulo de un test proyectivo.

Las técnicas proyectivas han sido desarrolladas principalmente dentro de un marco conceptual psicodinámico. Como se ha comentado, se considera que el material de estímulo, por su ambigüedad o baja estructura, provoca la proyección del mundo interno del sujeto y es este un concepto psicoanalítico, al menos en su origen, el cual será abordado con mayor amplitud posteriormente. En este punto, es útil referir alguna definición de personalidad para los propósitos de las técnicas proyectivas: en la teoría psicoanalítica o freudiana, se considera que las diferencias de la personalidad surgen de las diferencias individuales en la cantidad de energía psíquica (libido) disponible o en los métodos para manejar esta energía. (Lanyon y Goodstein, 1977)

También se han pretendido ubicar a las técnicas proyectivas con la teoría de la *Gestalt* y desde el paradigma E-R, entre otros. Algunas técnicas proyectivas que utilizan estímulos visuales se explican desde la teoría de la *Gestalt*, específicamente con el enfoque que plantea la relación de la personalidad con la percepción conocido como *New Look*. Weithemer (1957), Bruner (1948) o Gibson (1950 en Fernández-Ballesteros, 1992) consideran al Rorschach como perfectamente conceptualizable dentro de la psicología experimental gestaltista. Una de las hipótesis más importantes de Rorschach es la relación entre la percepción diferencial a los estímulos de la prueba y características específicas de la personalidad. Los conceptos como defensa perceptiva y distorsión perceptiva serán los mecanismos que permitan la integración personalidad-percepción a través de las técnicas proyectivas.

El modelo mediacional de Hull y el paradigma E-R puede ser utilizado como soporte de las técnicas proyectivas. Auld señala que los métodos proyectivos pretenden el análisis de los *hábitos emocionales*. Para explicar las relaciones entre los métodos proyectivos y los hábitos, Auld distingue tres tipos de situaciones: la situación de origen, la situación de test y la situación criterio. La relación entre las tres situaciones se debe a los hábitos emocionales que fueron aprendidos en la situación de origen, que aparecen en la situación del test y, por el poder de éstos de explicar la conducta, permiten predecir el comportamiento en la situación criterio. (Fernández-Ballesteros, 1992)

Lindzey (1961 en Rabin, 1968) realizó un importante estudio para revisar las relaciones entre las teorías psicológicas y las técnicas proyectivas, encontró muchos elementos de congruencia entre estos instrumentos y la teoría psicoanalítica. Otro intento importante es la obra de Rabin (1968) como editor en la que compila las investigaciones en torno a diferentes enfoques teóricos y las técnicas proyectivas como el Psicoanálisis, la Teoría del aprendizaje, la Teoría cognitiva, la Teoría del Nivel de Adaptación y la Teoría del Campo, entre otras. (Rabin ed. 1968)

SUPUESTOS BASICOS DE LAS TECNICAS PROYECTIVAS

Independientemente de la teoría con que se pretenda explicar o fundamentar estas técnicas, los usuarios de éstas concuerdan en los siguientes supuestos, según Fernández-Ballesteros, (1992):.

1. Se parte del supuesto de que el sujeto cuenta con una estructura básica y estable de la personalidad. Las respuestas de los sujetos ante las técnicas proyectivas permitirá la exploración de tal estructura.
2. Existen distintos niveles en tal estructura de la personalidad que son abordados por las diferentes técnicas. El Rorschach permite un estudio más profundo que el TAT por ejemplo.
3. Existe una relación entre los "inobservables" que integran la estructura de la personalidad y las manifestaciones conductuales.

Por su lado, Murnstein, 1961, (en Megargee, 1980) refiere otro conjunto de supuestos sobre las pruebas proyectivas que analiza cuidadosamente presentando evidencia de investigación que los apoyan o los contradicen. Su estudio abarca supuestos relativos a la ambigüedad del estímulo; la semejanza del estímulo al sujeto para aumentar la proyección; la poca importancia del estímulo, dada su ambigüedad, en comparación con el valor de la respuesta emitida; la significación de cualquier respuesta a las pruebas para el análisis de la personalidad; la complementareidad de varias pruebas proyectivas para explorar diversas capas de la personalidad; la falta de conocimiento del sujeto de lo que descubre de sí mismo en una prueba proyectiva, entre otras.

EL CONCEPTO DE PROYECCION

Las definiciones de los tests proyectivos implican dos puntos importantes: el concepto de proyección y el de estímulo proyectivo.

El estímulo proyectivo. Tanto el Rorschach como el T.A.T., las dos técnicas proyectivas más usadas, utilizan estímulos visuales ambiguos que el sujeto tiene que describir. El Rorschach, el test que para mucha gente es símbolo de la evaluación psicológica, utiliza manchas simétricas de tinta. El T.A.T. utiliza figuras humanas ambiguas. El término "proyectivo" se utiliza porque, dado que un estímulo es vago, se afirma que a menudo es tan vago que hace imposible una descripción verídica y el sujeto debe proyectar algo de sí mismo ante tal estímulo. (Kline, 1985, 1993) Como se verá más tarde, no todas las técnicas proyectivas utilizan estímulos ambiguos y el grado de ambigüedad es variable entre ellas.

El significado de la proyección. Según Kline, (1993) el mecanismo de defensa de la proyección no está necesariamente involucrado en las pruebas proyectivas aunque esto pueda ocurrir en algunos individuos y bajo ciertas circunstancias.

Bellak (1990) refiere que la proyección ha sido entendida por la psicología norteamericana como un mecanismo de defensa al servicio del yo, destinado a evitar la conciencia de los deseos, pensamientos e impulsos inaceptables, y a cumplir con su tarea asignando dichos fenómenos subjetivos indeseables al mundo objetivo. Este fue el concepto original de Freud pero posteriormente el mismo Freud lo amplió a un proceso perceptual general por lo que " toda percepción significativa actual esta fundamentada y organizada en las huellas mnémicas de todas las percepciones previas". A esto Bellak le llamó "distorsión aperceptiva" y agrega que podría ser considerado como el marco de referencia más útil para algunos de los llamados métodos proyectivos.

Las técnicas proyectivas, al igual que el psicoanálisis original, se han basado en una psicología del ello, en relación con los impulsos y su expresión. Bellak (1990) analiza el papel del yo en la producción imaginativa a la luz del concepto psicoanalítico de la naturaleza y las funciones del yo (formuladas por Hartmann) y considera que las pruebas proyectivas necesitan una nueva exploración a la luz de estas funciones del yo. Agrega que el monto de la participación del yo en las distintas producciones y el saber cual función del yo es la que interviene, varía de método a método y de caso a caso.

La participación del yo en las producciones imaginativas se ubica en un continuo que va de la mínima a la máxima participación, a saber:

Participación del Yo en las producciones imaginativas.

(min)	→sueño→	fenómeno hipnagógico→	fantasía preconscious→	ensañación→	asociación libre→	producciones artísticas→	conducta de prueba en técnicas proyectivas→	solución de problemas→	(max)
-------	---------	-----------------------	------------------------	-------------	-------------------	--------------------------	---	------------------------	-------

(Esquema adaptación de M Fuentes a partir de Bellak 1990)

(La mínima participación del yo se presenta en el sueño y la máxima se presenta en la solución de problemas, esta como producción imaginativa pues se refiere al proceso en que se obtiene la solución a un problema pendiente dentro de la fase de sueño o en la fase entre la vigilia y el sueño)

Bellak (1990) analiza cada uno de estos fenómenos o procesos y al llegar a la conducta frente a las pruebas proyectivas habla de una función *oscilante* del yo. Pone como ejemplo una aplicación de T.A.T. en la que se le pide al sujeto que deje correr su mente en forma libre -se le induce a una autoexclusión del yo- pero después se le solicita que diga que está pasando, qué fue lo que condujo a tal situación, cuál será el final, esto es, se le pide que se apegue al estímulo y las instrucciones o que se ancle de nuevo al yo. El desempeño en el T.A.T. y en otras técnicas proyectivas presupone una función *oscilante* del yo. El paciente exageradamente rígido no podrá disminuir el control y sólo proporcionará datos descriptivos y pobres y en el extremo, el paciente con insuficiente control del yo, abandonará el estímulo y la tarea.

La habilidad para relacionar desde el nivel latente al manifiesto (conductual) -de los datos de la prueba proyectiva hasta la conducta real-, ha sido considerada como el problema crucial de la evaluación proyectiva. En algunas situaciones es necesario hacer este tipo de inferencias, de lo latente a lo manifiesto: para fines teóricos, experimentales, en situaciones clínicas como la predicción de la conducta criminal o suicida y en problemas como la selección de personal. (Bellak, 1990)

Es necesario considerar el hecho de que no deberían hacerse interpretaciones de pruebas proyectivas a ciegas. El análisis de los datos del inconsciente debe hacerse como complemento de los datos conductuales. Por otro lado, es necesario asimismo considerar al yo como la variable que interviene entre el nivel latente y el manifiesto. Por ejemplo, en el T.A.T., un estudio de las funciones del yo en relación a los impulsos nos permite observar si los impulsos agresivos se manifiestan o no. En opinión de Bellak (1990), es necesario que las pruebas proyectivas dejen de ser análisis de impulsos inconscientes y se conviertan en el estudio de esos impulsos y su interacción con el yo.

CLASIFICACION DE TECNICAS PROYECTIVAS

Semeonoff (1973 en Kline, 1985) realizó un sistema clasificatorio de los tests proyectivos a partir de una encuesta sobre éstos. Este investigador clasifica por un lado los estímulos proyectivos y por otro, el manejo de las respuestas proyectivas.

Los estímulos pueden ser:

- verbales
- visuales
- concretos
- de otras modalidades de los sentidos.

Las respuestas pueden ser (usadas para):

- descripción
- clasificación o de diagnóstico
- terapia

En su obra, Bellak (1990) refiere una clasificación de las técnicas proyectivas hecha con base en los procesos llevados a cabo por la persona para responder a ellas:

1. Métodos basados en el estudio del contenido. (Lo que el paciente dice) Ej. el T.A.T.
2. El estudio de aspectos expresivo-estructurales. (Cómo el sujeto dice o hace algo) Ej. el Rorschach y la grafología que pertenecen a niveles subsemánticos del funcionamiento mioneural.
3. Funciones gestalticas. (Cómo el sujeto integra la gestalt) Ej. el Bender-Gestalt y el Rorschach. El T.A.T. en algunos casos en que sólo se considera una parte de la lámina o se aleja totalmente del estímulo.
4. Imagen corporal o autoimagen. (Expresión de la propia imagen) Ej. El Dibujo de figura humana.
5. Métodos de preferencia. El Szondi está basado en un sistema de preferencias selectivas como indicadores de la personalidad.

Bellak (1990) agrega que cuando Frank acuñó el término de técnicas proyectivas, 1939, resultó apropiado para esa época pero con el tiempo se ha tornado confuso debido a que existen técnicas a las que difícilmente es aplicable por la diversidad de procesos inherentes a ellas.

Fernández-Ballesteros (1992) considera que las clasificaciones de este tipo de pruebas no exceden el simple ordenamiento de un material diverso aunque por su parte, presenta su propia clasificación lograda a partir de los estímulos utilizados por las diferentes técnicas proyectivas:

- Estructurales. El estímulo es material visual y el sujeto debe decir lo que ve o lo que le parece, como el Rorschach.
 - Temáticas. Material visual sobre el cual el sujeto narrará una historia, como el TAT.
 - Expresivas. No hay un estímulo sino la consigna de dibujar algo, como el Test HTP o el de la Figura Humana.
 - Constructivas. Se le pide al sujeto construir algo a partir de unos objetos que se le proporcionan. Ej. El test del pueblo
 - Asociativas. El sujeto deberá manifestar verbalmente asociaciones frente a palabras, frases o cuentos presentados en forma oral o escrita.
- Una variante de éstas últimas pueden ser los Tests proyectivo-auditivos, como el *Sound apperception Test* (Bean, 1965) en el cual ciertos sonidos tienen que ser interpretados por el sujeto. (referido por Kline, 1993)

DESCRIPCION BREVE DE ALGUNAS TECNICAS PROYECTIVAS

1. El test de Rorschach. (1921)

Elaborada por el psiquiatra suizo Hermann Rorschach. Es el más conocido de los tests psicológicos. Consiste en diez láminas con manchas de tinta simétricas. A los sujetos se les solicita que observen y describan cada una. La puntuación y la interpretación son altamente especializadas. Dispone de más de cinco mil referencias en trabajos y tesis. Beck (1952) y Klopfer (1956) en Fernández-Ballesteros, (1992) han desarrollado sistemas detallados de puntuación para el Rorschach. Holtzman (1968, en Aiken, 1996) construyó una prueba de manchas de tinta, HIT, que se califica con respuestas de opción múltiple, esta prueba cubre muchos de los estándares psicométricos lo que aumenta su confiabilidad pero, al parecer pierde la riqueza de las respuestas libres a los estímulos. Esto hace suponer que al ganar en confiabilidad se pierde en validez.

2. El test de Apercepción Temática (Murray, 1958)

Los estímulos son una serie de láminas con representaciones gráficas ambiguas principalmente acerca de seres humanos. A los sujetos se les solicita hacer una historia de cada lámina, y usualmente se les muestran 20 de las 30 de la prueba, elegidas a juicio del aplicador. La interpretación de estas historias es un proceso bastante subjetivo e impresionista que se concentra en el análisis de las necesidades y personalidad del personaje principal de las historias. Se han adaptado procedimientos de calificación más sistemáticos y normalizados. (Bellak, 1986 en Aiken, 1996) Se han construido algunos esquemas de puntuación especialmente fiables como el de Zubin y sus colaboradores. (1965 en Lanyon y Goodstein, 1977) El sistema de McClelland (1953 en Lanyon y Goodstein) se enfocó a la evaluación de la necesidad de logro en el T.A.T. y obtuvo amplia difusión. Las variables como la edad, el sexo, el aspecto, la familiaridad del examinador condicionan la producción de las historias en el sujeto (Vickes, 1956. Tszdzuki y Kuze, 1957 en Fernández-Ballesteros, 1992)

3. El test de la Casa, Arbol y Persona, *HTP* (Buck, 1948 en Kline, 1985) y el test de Dibujo de una Figura Humana.

En el primero, se les pide a los sujetos dibujar una casa, un árbol y una persona y responder a preguntas acerca de los dibujos. El mismo autor ha recopilado investigaciones clínicas acerca de este instrumento.

Existen varias versiones del test de la Figura Humana, la primera y la más conocida es la de Karen Machover (1948 en Fernández-Ballesteros, 1992). En éste, se le pide al sujeto que dibuje una persona lo más completa posible. Posteriormente se le solicita que dibuje una persona del sexo contrario a la primera que dibujó y que complete figuras incompletas. La mayor parte de la obra de Machover está dedicada a las reglas de interpretación que entrañan los aspectos cualitativos de los dibujos. Estas reglas fueron derivadas aparentemente de la experiencia clínica de la autora. La interpretación de Machover de los dibujos es eminentemente psicoanalítica. Swensen (1957, 1968) y Roback (1968) citados por Lanyon y Goodstein, 1977, concluyeron que había poco respaldo para las interpretaciones o hipótesis de Machover, y que de hecho eran más los datos que contradecían directamente sus argumentos pero por otro lado, encontraron que la calidad general del dibujo se relaciona con un nivel amplio de adaptación del sujeto. Lewinson (1965 en Lanyon y Goodstein, 1977) llegó a la misma conclusión. Probablemente el dibujo de la figura humana sea una técnica valiosa pero no se ha desarrollado una forma consistente y válida de calificar.

La hipótesis proyectiva de éstos y otros tests de dibujos sustenta que si la conducta proyectada es una función de la personalidad globalmente considerada, al carecer de estímulo previo sobre el cual proyectarla, los resultados -en este caso los dibujos- tienen que estar en función de tal estructura interna, con una alta probabilidad. (Hammes, 1969 en Fernández-Ballesteros, 1992) El interés del dibujo como expresión psicológica en su alternativa interpretativa, asume los planteamientos psicoanalíticos que parten del supuesto proyectivo pero con frecuencia este tipo de tests carece de sustento empírico. Caligor (1970 en Fernández-Ballesteros, 1992) quien también desarrolló un test de la figura humana, refiere que se acepta la evidencia de que la realización del dibujo refleja tanto los progresos en la evolución como determinadas características del desarrollo integral implicadas en la conducta perceptiva y en las habilidades expresivo-representativas.

El test de la figura humana es ampliamente utilizado en México para fines de evaluación laboral pero con frecuencia se realiza una interpretación intuitiva, sin conocimiento de algún método interpretativo.

4. Los tests de Completamiento de Frases.

Parten de la presentación de un estímulo verbal, oral o escrito, de frases incompletas a las que el sujeto deberá responder con sus asociaciones, terminando la oración. Existen varias versiones de autores como Bloom, Rotter (Aiken, 1996), Kelly y Fisher, 1951 y una diseñada para aplicación laboral de Bruce, 1963. (en Fernández-Ballesteros, 1992) Estos instrumentos tienen la limitación de la falta de sistemas objetivos de interpretación.

Las pruebas de frases incompletas son bastante usadas en la evaluación laboral en México pero con frecuencia la interpretación se realiza sin el respaldo de alguna metodología o autor.

LA CONFIABILIDAD DE LAS TÉCNICAS PROYECTIVAS.

Los problemas especiales asociados con la determinación de las confiabilidades de algunas pruebas proyectivas han llevado frecuentemente a la conclusión inapropiada de que las consideraciones de la confiabilidad no deben aplicarse a estas

Se ha sostenido que los métodos usuales para evaluar la confiabilidad no son aplicables a los instrumentos proyectivos. El método de estabilidad temporal no es aplicable porque una reaplicación significa una experiencia psicológica diferente de la aplicación original y las técnicas proyectivas son especialmente sensibles a los cambios ligeros en el sujeto. El método de división por mitades es inapropiado para la mayoría de las pruebas más conocidas, por ej. el Rorschach o el TAT, por la misma estructura de la prueba. Los usuarios y defensores de estas pruebas aducen que los cambios en los resultados observados en el tiempo, por ej. en una reaplicación, corresponden más bien a cambios reales en el individuo más que a deficiencias en la consistencia del instrumento. (Lanyon y Goodstein, 1977)

Uno de los problemas al abordar la confiabilidad radica en el hecho de que no se considero ésta al construir la prueba y, en algunos casos, los sistemas de puntuación fueron desarrollados posteriormente a la elección de los estímulos y el diseño de la prueba. La versión de Holtzman de la prueba de manchas de tinta, que tuvo este cuidado desde el principio, arroja datos satisfactorios de su confiabilidad. (Lanyon y Goodstein, 1977)

La falta de estandarización de las instrucciones es otro problema que provoca que el examinador pueda influir significativamente en las respuestas del sujeto aunque él mismo no se percate de ello. Asimismo, la subjetividad implicada en la mayoría de las puntuaciones derivadas de los instrumentos proyectivos es otro problema que afecta la confiabilidad interjuez o intrajuez pero la inconfiabilidad de la calificación contribuye a la inconfiabilidad general de la prueba. Los diferentes sistemas de calificación del TAT arrojan confiabilidades intercalificador muy distintas: mientras Murstein (1963) refiere una confiabilidad de .74; Feld y Smith (1958) encontraron confiabilidades de .89 con el sistema de calificación ideado por McClelland y col. (Lanyon y Goodstein, 1977)

La consistencia intercultural o transferibilidad planteada por Catell (1964 en Lanyon y Goodstein, 1977 es especialmente importante al intentar usar una prueba en una cultura diferente a la que fue dirigida. Este tipo de consistencia es la que se desea resaltar por su importancia global y el impacto específico en nuestra cultura, nuestro idioma y nuestro papel de usuarios de pruebas construidas en el extranjero. Desafortunadamente, la metodología propuesta por Catell para estimar la consistencia intercultural no es aplicable a las técnicas proyectivas

La confiabilidad de la interpretación global.

En relación a las técnicas proyectivas, es más útil el concentrarse en la confiabilidad de la interpretación global (Lanyon y Goodstein, 1977) para lo cual existen técnicas apropiadas. Se han desarrollado métodos para calcular la confiabilidad de la interpretación global de las pruebas proyectivas.

Un procedimiento es el emplear una serie de escalas o dimensiones estimativas relevantes para la prueba y para la clase de información descriptiva requerida. La interpretación se realiza haciendo que el especialista asigne puntuaciones o posiciones en estas escalas a los sujetos. Un procedimiento parecido es la técnica de selección Q: un conjunto de 100 tarjetas con enunciados descriptivos de la personalidad que los examinadores utilizarán para calificar los resultados de la prueba proyectiva. Este método logra convertir una interpretación cualitativa en un puntaje numérico que permita la comparación.

El uso del análisis G

Se ha desarrollado un método para obtener puntuaciones y análisis fiables para enfrentar la objeción de la baja confiabilidad de la puntuación e interpretación de las pruebas proyectivas. Este es el Análisis G, desarrollado por Holley, (1973) y colaboradores, en especial Vergelius (1976). Este análisis se centra más bien en el método de puntuar las respuestas y de correlacionar a los sujetos. La siguiente es la descripción de Kline (1985)

Empieza por puntuar objetivamente las respuestas por la presencia o ausencia de variables, de una forma dicotómica. Posteriormente se correlacionan a los sujetos, se factorizan estas correlaciones y se forman los grupos de análisis del factor Q. Finalmente se encuentran aquellas variables que discriminen entre los grupos. Los sujetos, no las variables, son correlacionados utilizando el Índice G de acuerdo. Al factorizar a los sujetos, los factores resultantes saturan sobre ciertas personas que así forman un grupo.

El método abarca tantas variables como lo requieran las respuestas estudiadas. Si la respuesta existe, se califica con 1, si está ausente, se califica con 0. Usando este análisis G, Holley encontró una notable discriminación entre tres grupos: depresivos, esquizofrénicos y un grupo de control. Este método se ha seguido aplicando y perfeccionando.

Cualquier tests proyectivo puede calificarse así, incluso las entrevistas o los juicios intuitivos de los clínicos. El análisis G permite que los tests proyectivos sean utilizados en el estudio cuantitativo de la personalidad y abre amplias perspectivas al estudio de éstos.

La clase de confiabilidad establecida para cualquier instrumento de evaluación debe ser apropiada para el uso que va a darse al procedimiento. Si lo que se requiere es hacer predicciones acerca de acontecimientos especiales, la que debe examinarse es la confiabilidad de tales predicciones. Si lo que se busca es la interpretación global, la confiabilidad de esas interpretaciones es lo que debe investigarse. (Kline, 1985)

LA VALIDEZ DE LAS TÉCNICAS PROYECTIVAS

La libertad que las técnicas proyectivas permiten a las respuestas de los sujetos, que puede considerarse una de las principales virtudes de este tipo de instrumentos, es también su mayor debilidad porque origina la mayoría de los problemas al estimar su validación. (Megargee, 1980)

Los diferentes conceptos de validez y las formas de calcularla, desarrollados sobre todo para instrumentos psicométricos, entrañan dificultades y obstáculos importantes al tratar de aplicarlos a las técnicas proyectivas. Según la Asociación Psicológica Americana, la validez de un procedimiento significa el grado en que este es capaz de alcanzar sus metas. Con este razonamiento, si las pruebas difieren en sus metas, un enfoque que sería apropiado para demostrar la validez en una prueba pudiera ser inapropiado para otro (Lanyon y Goodstein, 1977)

Un artículo clásico de Ainsworth (1951 en Megargee, 1980) sobre el tema indica que "el proceso de validación de las técnicas proyectivas es más parecido al proceso científico común de validación de hipótesis que al de validación de un test". La técnica proyectiva trata con n funciones o variables e intenta describir a los individuos en términos de un patrón dinámico de funciones o variables interrelacionadas. La multiplicidad de variables presenta el problema más grande de validación. Otro problema es el uso apropiado del test en la investigación de validación. Discute asimismo el problema de la confiabilidad de las técnicas proyectivas, si un test no es confiable no puede ser válido y en esta afirmación está implícita la idea de que la función que se mide es una función estable. Finaliza con una exposición sobre los criterios externos contra los cuales comparar las hipótesis interpretativas.

La presencia de variables extrañas en el proceso de evaluación con tests proyectivos afectan los resultados: Vernon (1964 en Kline, 1985) identificó la actitud del respondiente y del aplicador, el sexo o la raza del examinador, etc. Esta susceptibilidad a la situación resulta rara si en verdad los tests proyectivos miden las capas más profundas de la personalidad.

En referencia al método propuesto por Eysenck (1959 en Kline, 1985) para investigar la validez de los tests proyectivos que plantea la interpretación de los resultados "a ciegas", debe considerarse que por la misma naturaleza de este tipo de tests no es posible interpretar los resultados fuera del contexto de la vida del propio sujeto.

Un concepto valioso, útil para las pruebas de personalidad proyectivas o no, es el de validez incremental: es el grado en que una prueba mejora la precisión de la predicción sobre el nivel que se alcanzaría sin la prueba. Las pruebas del tipo de Rorschach frecuentemente son interpretadas como parte de un conjunto de datos que abarcan entrevistas, historiales, etc. Este tipo de validez se refiere a qué tanto la prueba aumenta la eficiencia predictiva sobre la información que de otra forma se podría obtener fácilmente. (Lanyon y Goodstein, 1977) Para mayor información y estudios acerca de la validez incremental se puede consultar a Groth-Marnat, (1990) p 20

Se ha desarrollado otro tipo de validez para los tests utilizados en la práctica clínica: la validez conceptual (Maloney & Ward, 1976, en Groth-Marnat, 1990). En contraste con los métodos usuales de validez que se enfocan al instrumento en sí, la validez conceptual se enfoca a los individuos con sus historias y conductas particulares.

La validez conceptual es un medio de evaluar e integrar los datos de los tests para que las conclusiones de los clínicos formen postulados precisos acerca del examinado. La validez conceptual también tiene que ver con los constructos evaluados, como en la validez de constructo, pero en este caso los constructos se refieren al sujeto más que a la prueba en sí.

La validez de concepto produce constructos como un producto final después de una serie de procedimientos que abarcan tests, observación conductual, historia personal, etc. Dichos constructos resultantes son útiles para proveer información que pueda ser usada para ayudar a resolver los problemas particulares que una persona pueda estar enfrentando (Groth-Marnat, 1990).

Como se ha visto con esta breve exposición, la cuestión de validar las técnicas proyectivas implica problemas muy diferentes a los que presentan los instrumentos psicométricos y el camino está abierto a los psicólogos para diseñar procedimientos especiales para estimar la validez y la confiabilidad de dichas técnicas.

DEFENSA DE LOS TESTS PROYECTIVOS

Kline (1993) plantea algunas de las razones para continuar con el uso de los tests proyectivos.

1. Proveen información única que no se puede obtener por otro medio.
2. Algunos resultados de los tests proyectivos sugieren que estos pueden ser técnicas poderosas. En manos de un evaluador experimentado se pueden obtener resultados valiosos.
3. La riqueza de los datos de los tests proyectivos. De ellos se obtiene información fascinante que parece profunda e importante para entender la personalidad.
4. El éxito de algunos métodos objetivos de calificación desarrollados para algunas técnicas proyectivas, especialmente del Rorschach y del TAT puede convertirlos en valiosos instrumentos de evaluación de la personalidad.

APLICACION DE LAS TECNICAS PROYECTIVAS EN AMBITOS LABORALES

Las técnicas proyectivas siguen siendo usadas en escenarios laborales a juzgar por la gran cantidad de artículos publicados sobre el tema. Buross (1978 en Fernández-Ballesteros, 1992) refiere que el Rorschach y el TAT aparecen como la segunda y tercera técnica en cuanto a referencias bibliográficas. Un estudio de Piotrowsky et al., 1985 (en Aiken, 1996) encontró que estas técnicas ocupan el segundo y cuarto lugar en cuanto a su uso entre miembros de la Sociedad para la Evaluación de la personalidad. En una encuesta realizada en empresas de México en 1994, el test de la figura humana y el HTP figuran entre los cuatro instrumentos más utilizados para la evaluación de la personalidad (Milán, en Arias Galicia, 1996).

¿Técnicas proyectivas o pruebas psicométricas?

La disyuntiva no tiene por qué plantearse en términos de elegir los instrumentos psicométricos o las técnicas proyectivas en la evaluación de personal. Los defensores de una o de otra aproximación con frecuencia no están preparados en ambas. Tanto el enfoque proyectivo como el psicométrico tienen sus ventajas y sus desventajas y en ciertos casos pueden combinarse con éxito si el usuario conoce bien los instrumentos que planea combinar.

Levy (1963 en Lanyon, 1977) entre otros psicólogos, considera que la distinción entre los instrumentos proyectivos y los inventarios es falsa debido a que los inventarios pueden ser contestados en forma "proyectiva" y que los perfiles resultantes pueden ser utilizados por clínicos diestros como base para una interpretación global de la personalidad. Agrega que las técnicas proyectivas pueden ser calificadas objetivamente y luego interpretadas a la luz de datos normativos.

Subsiste la limitación en las técnicas proyectivas de la aplicación e interpretación individual, del nivel de especialización del psicólogo requerida para usarlas y del costo que implican ambos puntos que, por otro lado, se justificaría en el caso de evaluar personas-clave en una organización. Lo que es muy discutible es el empleo de instrumentos sin la debida preparación del psicólogo y la toma de decisiones organizacionales a partir de una interpretación así obtenida.

Las técnicas psicométricas han demostrado su utilidad en labores de selección rutinaria de grandes grupos de personas y depende de la creatividad (y la preparación del psicólogo) para utilizar y sacarle el mayor provecho a esos datos. Los inventarios de personalidad permiten comparaciones entre individuos, entre grupos, intragrupos, interculturales, interépocas, intraindividuos, etc. Facilitan los estudios de validación concurrente o predictiva con criterios reales de desempeño así como la comparación en el tiempo y en el espacio, pues es factible comparar grupos de diferentes culturas o individuos frente a esos grupos.

El proceso irreversible y acelerado de globalización que estamos viviendo en todas las áreas de la actividad humana obliga a mantener vigente las técnicas de evaluación. Si hubo épocas en que el proceso de evaluación era una experiencia local y temporal para cubrir determinados objetivos, actualmente se da el caso, por ejemplo en empresas transnacionales, que tenga que evaluarse un individuo o un grupo en un país para realizar ciertas funciones en otro país, por lo que el psicólogo requiere de creatividad y preparación por supuesto para que la interpretación de los resultados de la evaluación sea útil para el propósito específico.

El psicólogo decide (o debería decidir) qué instrumentos usar con base en el objetivo de la evaluación. Existe amplia variabilidad entre los psicólogos tanto en los procedimientos de recolección de información como en la integración de los resultados. El factor tiempo es una consideración importante, el factor económico es otro y el factor de su propia especialización es otro que debe tomarse en cuenta.

La consecuencia más importante del enfrentamiento entre estos dos enfoques ha sido la pérdida de colaboración para que ambos progresen. Con excepción del test de Holtzman, los adelantos que pudieran haber representado una síntesis de las dos tradiciones han sido insignificantes. Lanyon y Goodstein, (1977).

EL FUTURO DE LAS TECNICAS PROYECTIVAS.

Desde la perspectiva psicodinámica, las técnicas proyectivas son consideradas procedimientos de evaluación útiles y adecuados para sus fines y la continuidad de las investigaciones publicadas indican que, lejos de abandonarse, existe un repunte en su uso.

La revisión de los aspectos de confiabilidad y validez lleva necesariamente a:

1. desarrollar métodos especiales para calcular la confiabilidad y la validez de las técnicas proyectivas.
2. Reestructurar los instrumentos existentes, específicamente la forma de calificar para hacer los resultados más accesibles a la investigación.
3. Diseñar nuevas técnicas proyectivas que aprovechen los aspectos positivos y minimicen sus limitaciones y
4. Desarrollar técnicas a partir de teorías más recientes que abran perspectivas al ampliar su uso fuera del ámbito clínico. (Fernández-Ballesteros, 1992)

2.3 EL ENFOQUE PSICOMETRICO.

El Modelo del Rasgo descrito anteriormente permite ubicar claramente el enfoque psicométrico en la Evaluación Psicológica en cuanto a sus características generales. Quizá la característica más relevante, la que define a este enfoque, es la que le da su nombre: la cuantificación de los atributos personales.

El enfoque psicométrico se ubica dentro de la aproximación nomotética, cuantitativa y utiliza las estrategias correlacionales que se caracterizan por medir los rasgos de la personalidad y determinar "qué va con qué". La estrategia fcorrelacional trabaja con múltiples variables simultáneamente. La medición psicologica depende, en gran parte, de desarrollos estadísticos y las computadoras han apoyado el avance de la Psicometria al permitir manejar una gran cantidad de datos. (Buss y Poley, 1979)

Existe una amplia variedad de instrumentos psicometricos que se mencionarán sucintamente ya que el objetivo de este trabajo son los Inventarios Factoriales de personalidad, en especial el 16 PF de Catell.

La tradición psicométrica se desarrolla de manera importante por el interés de medir la inteligencia, sobre todo, y por la necesidad de evaluar una gran cantidad de personas en la Primera Guerra mundial. Una breve historia de la Psicometria servirá de introducción al tema.

BREVE RESEÑA HISTORICA

Desde el principio de la historia de la humanidad se ha reconocido el hecho de que las personas difieren en sus capacidades, personalidad y comportamiento y que estas diferencias pueden evaluarse de alguna forma.

La evaluación de las características individuales con un propósito particular se remonta a las civilizaciones antiguas. Existen referencias de que los antiguos chinos (1115 A.C.) utilizaban instrumentos de evaluación de escritura y aritmética para elegir al personal de cargos públicos. Los griegos también hicieron su aportación, con Platón y Aristóteles, quienes escribieron acerca de formas de seleccionar personas para ocupar puestos en el ejército.

Durante siglos se consideró al individuo como una unidad indivisible lo que frenó el estudio de los rasgos o características por separado. Durante la Edad Media, el interés por las diferencias individuales era casi inexistente. En la estructura social de la sociedad europea medieval, las actividades de una persona se determinaban por la clase social a la que pertenecía. Para el siglo XVI, creció la idea de que las personas eran únicas y estaban capacitadas para hacer valer sus dotes naturales. El Renacimiento reestablece el individualismo (Buss y Poley, 1976)

Fué hasta el siglo XIX, en la Gran Bretaña, con el desarrollo del capitalismo, que surge la necesidad de una sociedad laboral especializada. Los individuos que presentarían las mejores características tendrían las mejores oportunidades de trabajo.

"Desde el punto de vista sociológico, puede considerarse que este enfoque individual se debe, principalmente a las fuerzas complementarias del protestantismo y del capitalismo...una sociedad capitalista creaba diferencias individuales en el sentido que proporcionaba oportunidades para una especialización del talento humano, que no había tenido precedentes...La aparición de una economía capitalista estaba produciendo una nueva imagen del hombre, que preparaba el camino para la medición y la cuantificación de las características psicológicas..." (Buss y Poley, 1976)

Los escritos de Charles Darwin sobre el origen de las especies estimularon el estudio de las diferencias individuales y el surgimiento de la Psicología científica, cuyo origen se ubica en Alemania a fines del siglo XIX. Darwin se interesó en las bases heredadas de la inteligencia y en las técnicas para medir las capacidades.

Los psicólogos alemanes demostraron que los fenómenos psicológicos se podían expresar en términos cuantitativos y racionales. Los primeros psicólogos experimentales comenzaron a explorar las capacidades mentales utilizando técnicas cuantitativas, como los métodos psicofísicos de Weber y Fechner para medir la agudeza visual, la auditiva y la del tacto. En otra vertiente, en Francia, los psicólogos y psiquiatras que estudiaron las enfermedades mentales influyeron en el desarrollo de las pruebas y técnicas de evaluación clínica. En Estados Unidos, la atención que se le dio a los exámenes educativos dió como resultado el desarrollo de medidas estandarizadas de logro académico (Aiken, 1996)

Francis Galton, James Cattell y Alfred Binet tienen una especial importancia como pioneros de la naciente Psicometría. Galton elaboró varias pruebas sensoriomotrices y diseñó varios métodos para investigar las diferencias individuales en las capacidades y el temperamento (Aiken, 1996). Quizá sus contribuciones más importantes sean: el descubrimiento de que muchos atributos humanos tienden a distribuirse según la curva normal y el desarrollo de métodos de correlación para comparar dos conjuntos de mediciones.

J.M.Cattell, con conocimientos e influencia tanto de alemanes como de los ingleses, retorna a Estados Unidos y se dedica a investigar la relación entre medidas psicofisiológicas y el logro académico. El término "test mental" fue utilizado por primera vez por este psicólogo para designar instrumentos con los que evaluaba las diferencias de algunos atributos entre estudiantes universitarios.

Es Binet, en Francia, quien elabora la primera prueba mental que probó ser un indicador efectivo del logro académico. La Escala de Binet-Simon fue también el primer intento de proporcionar una escala numérica de niveles de inteligencia que posteriormente se denominaron Edades Mentales. Esta escala fue ampliamente usada, estudiada y posteriormente, al revisarla, se conoció como la prueba Terman-Merrill (Aiken, 1996)

Otros pioneros de la evaluación y de los tests psicológicos son: Spearman con su Teoría de los Tests (Factor G), Thorndike con sus tests de capacidad educativa, Terman con tests de inteligencia, quien introdujo en concepto de C.I., Woodworth y Rorschach en los tests de personalidad; Strong en la medición de los intereses. El trabajo de Otis en los tests de inteligencia que se aplican en grupo llevo de manera directa a la elaboración de los exámenes Alfa y Beta del ejército de Estados Unidos, durante la Primera Guerra. (Aiken, 1996)

A partir de la Primera Guerra, muchos estudiosos han contribuido a la teoría y práctica de los tests psicológicos y educativos no sólo en la concepción y diseño de las pruebas sino en el progreso en la metodología estadística y los avances tecnológicos en la preparación y calificación de los tests así como en el análisis de los resultados. (Aiken, 1996)

Esta tendencia de elaboración científica de pruebas psicológicas progresó notablemente con Thurstone hacia 1930 con sus estudios sobre análisis de factores. La importancia de Thurstone radica en el desarrollo de una herramienta estadística por excelencia para la identificación de los multirrasgos, el análisis factorial y los métodos multivariados. Con esa metodología, Thurstone identifica siete habilidades primarias (Buss y Poley, 1976)

El interés por el diseño y la utilización de pruebas creció y surgieron, entre otras, las pruebas de personalidad que fueron las que abrieron realmente el mercado por la popularidad de que gozaron. A partir de ahí, innumerables investigadores se han dedicado a este campo y pueden contarse por miles las pruebas existentes.

En la década de los 60's surgió en Estados Unidos principalmente, un movimiento en contra de la utilización de las pruebas de inteligencia y posteriormente de otras pruebas psicológicas sobre la base de la discriminación por educación, orígenes familiares, raza, sexo, etc. A partir de 1970, varios estados han promulgado legislaciones para prohibir el uso de las pruebas de C.I. en las escuelas, sobre la base de que tienen desviaciones culturales y que no miden con precisión la inteligencia. La Asociación de Psicólogos Negros pidió una moratoria completa sobre las pruebas, hasta que se conociera algo más respecto a que miden las pruebas de C.I. (Vernon, 1982) pero éstas y otras legislaciones no han desalentado el uso y la investigación de las pruebas. En fechas recientes se han publicado revisiones de dos de los más importantes tests psicométricos: del WAIS y del MMPI. El apoyo de la computación amplia enormemente las aplicaciones de la Psicometría. Las normas legales que restringen el uso de las pruebas necesariamente alientan a los psicólogos a construir instrumentos cada vez mejores.

Con la convicción de que las personas son diferentes entre si en un gran número de características y que esas diferencias pueden ser medidas comienza el estudio de las diferencias individuales y así nace la Psicometría.

BASES TEORICAS DEL ENFOQUE PSICOMETRICO

La principal base teórica de este enfoque es la Psicología Diferencial, abordada desde la perspectiva de la teoría de los rasgos. El enfoque teórico condiciona la construcción de los tests en la mayoría de los casos: las pruebas que se ubican dentro del Modelo del Rasgo se construyen con criterios racionales, empíricos y factoriales.

Algunos de los supuestos acerca de la inteligencia referidos por Vernon pueden generalizarse como supuestos del enfoque psicométrico: Los rasgos son atributos reconocibles que ocasionan diferencias entre las personas, son entidades homogéneas que pueden variar en cantidad pero que son esencialmente estables en su naturaleza durante toda la vida. El muestreo de tareas apropiadas y la normalización de las puntuaciones en comparación con la distribución en la población general, proporciona medidas aceptables como medidas cuantitativas. (Vernon, 1982)

Los tests psicológicos son los principales instrumentos de medición representantes de este enfoque. Cronbach (1990 citado por Fernández-Ballesteros, 1992) señala que un test es un procedimiento sistemático para observar la conducta y describirla con la ayuda de escalas numéricas o categorías establecidas. La Psicometría se define como la teoría e investigación con respecto a la medición de las características psicológicas (Aiken, 1996)

LA MEDICION EN PSICOLOGIA

El reconocimiento de la existencia de las diferencias individuales lleva a su estudio sistemático. Ya no es solamente la intención de clasificar o tipificar sobre la base de algún atributo sino es la preocupación por cuantificar esas diferencias.

"En la vida diaria, la palabra 'medición' tiene un significado claro y conciso. Para medir en situaciones prácticas generalmente disponemos de instrumentos que nos dan resultados fijos y precisos en forma de puntajes. La situación es diferente, sin embargo, cuando queremos medir variables psicológicas. Una variable psicológica se define aquí como un propiedad o característica que poseen diferentes individuos en cantidades distintas. Al medir variables como independencia, neuroticismo, capacidad de pensamiento lógico o aptitudes de aprendizaje, nos encontramos con problemas de escalamiento muy complejos. El problema de medición en Psicología es considerablemente más complicado." Medir es dar la magnitud de cierta propiedad de uno o más objetos con ayuda del sistema numérico. Los números usados de esta manera pueden llevar diferentes cantidades de información (Magnusson, 1975)

Sobre la misma idea, Nunnally (1987) agrega que la asignación de números a los objetos debe estar basado en un conjunto de reglas y persigue el fin de representar cantidades de atributos. Al mencionar reglas significa que el procedimiento para asignar números debe formularse explícitamente. Agrega que la medición en la ciencia tiene dos ventajas principales: permite una mayor precisión en la comunicación de los resultados que si sólo se usaran juicios personales y permite el empleo de métodos precisos de análisis matemático.

Magnusson (1975) plantea distinguir tres niveles de medida, los cuales difieren en la cantidad de información llevada por los números que representan las magnitudes de las cualidades. Los números pueden dar dichas magnitudes sobre una escala ordinal, una escala de intervalo o una escala de proporción, las que se describen a continuación junto con la escala nominal como referencia ya que esta no es propiamente una escala de medida.

-La ESCALA NOMINAL. Propiamente no implica medición, pues se asignan numerales con fines únicamente de diferenciación, clasificación o nominación, por ejemplo, en una escuela se dice para diferenciar los grupos de primer año: Grupo 1, Grupo 2, Grupo 3. No existe ninguna relación cuantitativa entre esos grupos, el grupo 2 no significa que sea mayor que el grupo 1.

-La ESCALA ORDINAL. Establece jerarquía con números. Se asignan números para indicar el orden entre ellos sobre el atributo medido. Por ejemplo, la Escala de Dureza de Mohs, La Escala sísmica de Richter.

En este tipo de escalas no existe un cero absoluto. No necesariamente existe la misma distancia entre número y número. Los números solamente proporcionan el orden de los objetos con respecto al rasgo que mide. Los modelos estadísticos aplicables son: modo, frecuencia, mediana, percentil, coeficiente de contingencia y de correlación.

-La ESCALA INTERVALAR. Se asignan numerales y existe la misma distancia entre número y número; cuenta con un cero relativo. En esta escala, los números dan información acerca del tamaño de las diferencias entre los objetos con respecto a la magnitud del rasgo medido. Las diferencias entre los números pueden compararse entre sí. Ejemplo, la Escala Farenheit o Centígrada para medir temperatura. Los modelos estadísticos aplicables son: Además de las anteriores, media aritmética, desviación estándar y coeficiente de correlación producto-momento de Pearson.

-La ESCALA DE RAZÓN O PROPORCIÓN. Cuentan con un cero absoluto y éste es el comienzo de la medición. Existe la misma distancia entre número y número. En esta escala, los números dan información, no sólo del orden de rango de los objetos y del tamaño relativo de las diferencias sino también de la relación entre las proporciones. Ejemplo: Sistema métrico decimal. En este tipo de escala es posible aplicar todos los estadísticos.

Los instrumentos de medición en Psicología utilizan escalas ordinales y en el mejor de los casos, intervalares. No se ha llegado a desarrollar instrumentos psicológicos que utilicen escalas de proporción por muchas limitantes entre ellas por el requisito de establecer un punto cero. ¿Existen casos en que algún atributo psicológico sea inexistente? (Magnusson, 1975)

No existe forma absoluta de medir algún atributo psicológico porque las únicas medidas con que contamos requieren de referencias, es decir, al medir algún atributo se requiere contar con las calificaciones de un grupo de estandarización que nos permita interpretar esa medida, por ejemplo, en una prueba de destreza manual la referencia o los parámetros sirven de base para determinar si el puntaje de un individuo es más alto o más bajo que el promedio del grupo. Magnusson (1975) lo expresa de la siguiente manera:

1. La posición de un individuo sobre un continuo no es dada como un puntaje absoluto sino como un puntaje relativo. Si se tiene solamente un individuo, su posición sobre un continuo psicológico no puede medirse pues solo podemos comparar a los individuos entre sí.

2. Los instrumentos para medir variables psicológicas sobre una escala de intervalo se construyen empezando por suponer una distribución normal de los puntajes. En esta escala de intervalo usamos la variación entre cada individuo como unidad de medida. La significación y exactitud de la escala obtenida depende, por supuesto, de lo correcto de la suposición

LA CURVA NORMAL

La mayor parte de los valores obtenidos de un grupo de personas -como es el caso de las pruebas psicológicas- se distribuyen en forma de curva normal. La distribución normal gráficamente se observa como una curva en forma de campana que se extiende a lo largo del eje x en ambos sentidos.

Las características de la distribución normal están completamente determinadas y la frecuencia de cada puntaje dado puede obtenerse directamente de una ecuación propia de la curva normal. La curva normal es teórica, simétrica y asintótica. Para caracterizar una distribución normal se necesita al menos una medida de tendencia central y una de dispersión. (Magnusson, 1975)

Medidas de tendencia central de la curva normal

Media aritmética: promedio aritmético de los valores.

Mediana: Valor que divide el total de los datos en dos partes iguales.

Modo: Valor más frecuente.

De dispersión.

Son indicadores del grado de diseminación de los datos. Una dispersión alta dará una curva amplia. Una dispersión baja, dará una curva reducida. Aunque existen muchas posibles medidas de variación, las más útiles por sus propiedades descriptivas son la varianza y la desviación estándar.

Varianza s^2 : Es la medida del grado de variabilidad de las distribuciones. Una prueba debe tener una varianza alta; esto significa que aun cuando se distribuya normalmente, las calificaciones deben presentarse en todos los valores de la variable que se está midiendo.

Es el promedio de las puntuaciones desviadas, elevado al cuadrado.

Su limitante es que expresa la dispersión en unidades distintas a la que mide dicha variable. Al obtener raíz cuadrada de esta medida, retornamos a las unidades originales y obtenemos la desviación estándar.

Desviación estándar s , que se expresa en las mismas unidades de la medida considerada. Esta medida es sumamente útil en el manejo de los puntajes. (Nunnally, 1987)

REQUISITOS DE LOS TESTS

Los tests deben cubrir requisitos de orden metodológico que garanticen sus resultados. Los más importantes son la confiabilidad, la validez, la estandarización y la objetividad las cuales se explican a continuación, junto con otras cualidades como la discriminación y la sensibilidad.

CONFIABILIDAD.

La prueba como instrumento de medición debe ser consistente. La confiabilidad es la consistencia de la medición. Es la exactitud de la medición. Mientras mayor es la diferencia entre medidas del mismo rasgo en repetidas ocasiones, menor será la confiabilidad del instrumento. Es la cualidad que hace que una prueba mantenga similares resultados aplicada dos veces a la misma persona en circunstancias idénticas. La confiabilidad indica si la prueba es constante en sus mediciones. La falta de confiabilidad se ilustra con la suposición de medir una superficie o un objeto con un metro elástico: todas las ocasiones en que se midiera, las medidas resultarían diferentes. La confiabilidad se ve afectada por el error variable en la medición. (Magnusson, 1975)

La confiabilidad se mide por varios métodos:

a) Estabilidad temporal. (test-retest) Se aplica la misma prueba a los mismos sujetos después de cierto tiempo; la correlación obtenida se llama Índice de estabilidad temporal.

b) Homogeneidad o consistencia interna. Es un método muy empleado por su facilidad. Consiste en dividir la prueba en dos partes, por ejemplo, preguntas pares e impares, y en calcular la correlación entre los resultados obtenidos en las dos mitades aplicadas a los mismos sujetos. Este método no puede emplearse en pruebas con límite de tiempo; para hacerlo, se requiere de un método más complicado.

c) Equivalencia. Este método consiste en elaborar pruebas paralelas, es decir, pruebas similares que comprenden preguntas de naturaleza y dificultad análoga. Se calcula la correlación entre los resultados de los mismos sujetos en ambas pruebas. Estas pruebas se construyen al mismo tiempo. Esta forma es particularmente útil en la evaluación de la personalidad para verificar los resultados o en los casos en que el evaluado manipule el instrumento. Son pocas las pruebas que cuentan con formas paralelas; un ejemplo es el 16 PF de Cattell, que cuenta con formas equivalentes A y B y puede obtenerse un coeficiente de formas paralelas.

d) Confiabilidad entre calificadoros. Las calificaciones que asignan dos calificadoros distintos de una prueba objetiva son idénticas, en teoría. Es posible calcular este acuerdo. La confiabilidad entre calificadoros es el grado de concordancia entre los puntajes de dos o más calificadoros. (Aiken, 1996)

Otros dos métodos valiosos para estudiar la confiabilidad son: el método de análisis factorial y el del análisis de varianza. El primero (Guilford 1956 en Kline, 1985) se da cuando un test es sometido al análisis factorial junto con otros tests entonces los cuadrados de sus saturaciones pueden utilizarse para demostrar su confiabilidad. Se trata de un procedimiento muy potente pero depende de la adecuación técnica del análisis factorial en sí y de los tests con los cuales se haya factorizado la prueba. El segundo (Hoyt, 1941 en Kiine, 1985) plantea que la varianza del test queda repartida en dos fuentes: ítems y sujetos, (junto con su interacción, por supuesto) Cuanto menor sea la varianza debida a los ítems, tanto mayor será la confiabilidad del test.

Catell (1964 en Lanyon y Goodstein, 1977) propuso otra forma de medir la consistencia de una prueba que se considera de especial importancia porque se refiere a la estabilidad de la medida de los rasgos de personalidad entre diferentes culturas: la transferibilidad. Se refiere a la concordancia en cuanto al significado de las puntuaciones de la misma prueba aplicada al mismo tiempo a diferentes grupos (o culturas). La medida de la transferibilidad de un instrumento entre diferentes poblaciones, culturas o generaciones permite hacer investigaciones de la estabilidad intercultural del rasgo y es importante para determinar las limitaciones de su uso.

Una prueba puede ser confiable sin ser válida, pero no puede ser válida si no es confiable. La confiabilidad es una condición necesaria pero no suficiente para la validez. (Aiken, 1996, p 95) Kline (1985) comenta que la importancia de la confiabilidad per se es pequeña y que el énfasis sobre ella deriva del hecho de que una alta validez generalmente requiere una elevada fiabilidad pero por otro lado, debe reconocerse que es muy difícil demostrar la validez de los tests de personalidad y que la mayoría de estos no son válidos como lo demuestran las revisiones de tests realizadas por Buros (1978) Kline finaliza diciendo que en la práctica, los tests válidos son sumamente consistentes. Esta cuestión toma un cariz totalmente diferente al aplicarlo a técnicas proyectivas como se expuso al tocar el tema.

Todo puntaje obtenido con una escala psicológica se considera como la suma de un puntaje verdadero más un error de medición.

$$X_1 = T_1 + E_1$$

Donde:

X_1 es el puntaje del individuo 1

T_1 es el puntaje VERDADERO del individuo 1

E_1 es el error de medida del individuo 1

En la teoría clásica de los tests se supone que la calificación observada que obtiene una persona en una prueba se compone de una calificación "real" más algún error no sistemático de medida. La calificación real se define como el promedio de las calificaciones que se obtendrían si una persona presentara la prueba una cantidad infinita de veces. Se debe enfatizar que la calificación real nunca puede medirse con exactitud sino que debe calcularse a partir de la calificación observada que obtuvo la persona en la prueba. (Aiken, 1996)

Al ser imposible determinar el valor absoluto del error, tampoco se puede establecer el valor absoluto del puntaje verdadero. El error de medida se presenta como un sesgo sistemático o como un error aleatorio. Se dice que una medida es confiable cuando el error aleatorio es pequeño. (Nunnally, 1987)

VALIDEZ

Es el requerimiento más importante de una prueba. Es la utilidad del instrumento. Es la respuesta a la pregunta ¿para qué sirve esta prueba? Es la cualidad que determina que una prueba mida precisamente aquello que pretende medir. Según la Asociación Psicológica Americana APA, la validez de un procedimiento significa el grado en que éste es capaz de alcanzar sus metas (Lanyon y Goodstein, 1977) aunque, según Nunnally, (1987) en sentido estricto no se valida un instrumento de medición sino el uso que se le da. Existen diferentes tipos de validez:

Validez de Contenido. Se refiere a que el contenido de la prueba sea representativo de lo que se pretende medir. Es importante sobre todo en las pruebas de rendimiento escolar. una prueba de Geografía debe contener preguntas de esta materia. Se evalúa a través de juicios de expertos.

En los inventarios de personalidad, actitudes, motivaciones, etc. este tipo de validez en ocasiones facilita el falseamiento de las respuestas. Al dar a conocer al sujeto lo que se está investigando se propicia que ofrezca respuestas deseadas

Validez concurrente. Determina el grado de relación entre los puntajes de dos pruebas o de una prueba con un criterio de desempeño

Validez predictiva. Determina el grado en que una prueba sirve como indicador de una conducta futura; indica el grado en que las predicciones hechas con base en los resultados de las pruebas quedan confirmadas posteriormente. La apreciación de la validez predictiva y concurrente supone la existencia de criterios objetivos con los cuales asociar los resultados de una prueba. Es valiosa en el área laboral, clínica y educativa. Es especialmente importante en instrumentos usados para selección de personal por su capacidad de dar a conocer información del candidato antes de ser contratado. El procedimiento para validar una prueba es asociar y correlacionar estadísticamente los puntajes obtenidos en la prueba en el momento de ingresar a la empresa con un criterio de desempeño obtenido unos meses después, cuando el empleado ya se encuentra en funciones.

Nunnally (1987) al referirse a la validez predictiva considera que no se utiliza el término predicción en un sentido gramatical sino más general, para referirse a las relaciones funcionales entre un instrumento y los eventos de antes, durante* y después de su aplicación. (* engloba aquí a la validez concurrente). Por ejemplo, un instrumento aplicado a un adulto puede usarse para hacer predicciones sobre los eventos que ocurrieron en su niñez (postdición).

Validez de constructo. Se usa más bien en investigación. Averigua las cualidades psicológicas que trata de calibrar una prueba y los resultados que de acuerdo con una teoría podrían esperarse. Se refiere al grado en que un instrumento es una medida de un constructo particular. Este tipo de validez es especialmente útil en los inventarios de personalidad y es la usual en inventarios factoriales (Aiken, 1996)

En México existen pocos estudios de validación de instrumentos pero los resultados tienen poca o nula difusión. Estos se realizan más en ámbitos universitarios que en las empresas.

OBJETIVIDAD.

La prueba debe contar con métodos de puntuación claros y no ambiguos (Colom, 1995), con criterios invariables que no permitan la injerencia del juicio del calificador. Es objetividad en contraposición con subjetividad. Las calificaciones que calculan dos calificadores distintos de una prueba objetiva serán idénticas. (Aiken, 1996)

Las pruebas objetivas siempre arrojarán resultados cuantitativos. Una ventaja adicional de las pruebas objetivas es que pueden ser calificadas en tiempos cortos y se requiere poco entrenamiento para hacerlo. Otra es que son factibles de aplicación grupal. Los programas computerizados de aplicación y calificación de pruebas eliminan casi completamente la subjetividad incluyendo los errores humanos involuntarios.

ESTANDARIZACION.

Este requisito tiene dos acepciones importantes.

Por un lado se refiere a que el procedimiento de aplicación y calificación deberá mantenerse uniforme. Las instrucciones asimismo deben mantenerse sin variación.

La otra acepción de la ESTANDARIZACION se refiere a que las pruebas deben aplicarse a una extensa muestra de individuos para obtener normas derivadas de una muestra representativa de la población a quien se destina la prueba. (Aiken, 1996) A este proceso también se le llama normalización.

El comportamiento registrado mediante una prueba es evaluado estadísticamente con respecto a un grupo de individuos. La clasificación del sujeto examinado en relación al grupo de referencia es la meta final de la prueba. Los resultados individuales deben ser comparados con aquéllos obtenidos en grupos bien definidos, semejantes al del individuo. En cada grupo se obtienen curvas de distribución de frecuencia, las cuales van a ser divididas en áreas según la preferencia del constructor de la prueba. Una persona se localizará en el área que le corresponda de acuerdo con su calificación.

El propósito principal del proceso de estandarización es determinar la distribución de calificaciones brutas en el grupo de estandarización. Los puntajes directos o brutos de la prueba son convertidos a puntajes normalizados que permiten su interpretación. Las normas deben resultar de un muestreo apropiado y representativo de la población-objetivo y los puntajes normalizados deben estar bien calibrados, expresados en puntuaciones manejables estadísticamente, con promedios y desviaciones estándar definidas. (Aiken, 1996 y Kline, 1985)

Para los inventarios de personalidad existen normas percentilares y normas de calificación estándar que representan las medidas en una escala de intervalo con la media y la desviación estándar que se desea. Entre estas, las más representativas son las calificaciones Z, las calificaciones T y las calificaciones T. Para pruebas de rendimiento están las CEEB. Para pruebas de Capacidad intelectual, el CI de desviación y las AGCT, entre otras.

Existen también las calificaciones estándar normalizadas que aplican un procedimiento para cambiar la distribución de las calificaciones brutas a calificación normal. Como ejemplo de éstas está la escala de estamina. (Aiken, 1996)

DISCRIMINACION

Otro de los requisitos de los tests es su capacidad para ser discriminativos. El poder discriminativo está en función de su longitud y la varianza que tiene como consecuencia. Una prueba debe tener una varianza alta, esto significa que aunque se distribuya normalmente, las calificaciones deben presentarse en todos los valores de la variable que se está midiendo. Kline (1985) establece que los inventarios de personalidad son razonablemente discriminativos, por su longitud y consiguiente varianza y superan a la mayoría de los métodos de evaluación de la personalidad.

Es posible obtener índices de discriminación de una prueba. Ferguson (1949 en Kline, 1985) ha desarrollado un método llamado delta d que varía en su valor como la correlación desde 1 hasta -1, siendo ésta la discriminación máxima posible contemplada por una distribución rectangular. Para una distribución normal, delta es de 0.92. Según Kline (1985) las exigencias de un buen test psicológico se resumen en confiabilidad, validez y discriminación.

SENSIBILIDAD

Un test debe ser capaz de detectar diferencias individuales por muy pequeñas que éstas sean, independientemente de que las muestras sean muy homogéneas. (Colom, 1995)

CLASIFICACION DE LOS TESTS

A continuación se presentaran varias clasificaciones de instrumentos de evaluación, con base en criterios diversos, enfatizando en aquellos enfocados a la personalidad.

La clasificación de Kline (1985) de instrumentos para la evaluación de la personalidad es la siguiente:

- a) los tests psicológicos:
 - los inventarios,
 - las técnicas proyectivas
 - los tests objetivos de Catell*.
- b) las escalas de valoración,
- c) las observaciones de comportamiento,
- d) los diferenciales semánticos,
- e) la técnica de rejilla.

Kline originalmente había considerado a la entrevista como un instrumento de evaluación de la personalidad pero posteriormente la elimina por considerarla ineficaz con respecto a esta medida ya que sus indicadores de confiabilidad y validez son bajos.

* Pruebas T, son instrumentos situacionales en que el propósito de dicha prueba se oculta al sujeto puesto que este da una respuesta de comportamiento en una situación determinada. (Catell y Kline, 1982) Por su parte, Cronbach (1976 en Kline, 1993) prefiere llamarlos tests de desempeño.

Otra clasificación de instrumentos para evaluar la personalidad, ésta de Lanyon y Goodstein, (1977) es la siguiente:

- a) Pruebas de personalidad:
 - inventarios
 - técnicas proyectivas.
- b) Métodos para obtener información factual específica:
 - entrevista estructurada.
 - métodos de datos biográficos (ej. los formatos de solicitud)
- c) Métodos Situacionales:
 - las muestras de conducta.
 - las técnicas múltiples como los simulacros *assessment center*.

Existe una amplia variedad de instrumentos que son variaciones de algunos de los mencionados, como las listas de verificación o que es difícil de categorizar, como las pruebas de comunicación no-verbal, el Test de colores de Luscher, etc.

A partir de la clasificación proporcionada por Aiken (1996), se presenta la siguiente que se considera amplia e ilustrativa aunque no exhaustiva

- * Según el propósito de evaluación: educativos,
 - laborales,
 - clínicos.
 - sociales, etc.
- * Según los aspectos que mide: Cognoscitivos, Afectivos o Psicomotrices.
 - a) Cognoscitivos. Intentan cuantificar los procesos y productos de la actividad mental logro: es el que evalúa el conocimiento de alguna materia académica u ocupación; se concentra en el comportamiento pasado de la persona, lo que ha aprendido.
 - aptitud: centra la atención en el comportamiento futuro, lo que la persona es capaz de aprender con una capacitación adecuada. (ej. tests de inteligencia y habilidades especiales.
 - b) Afectivos. Evalúan características no cognoscitivas de la personalidad.
 - intereses, actitudes
 - valores, motivos
 - rasgos. etc.
 - c) Psicomotrices. Evalúan respuestas psicofisiológicas como el tiempo de reacción, la agudeza visual, la respuesta galvánica de la piel, etc.
- *Según la forma de aplicación: Individuales, colectivos o autoaplicadas. (Actualmente, con el desarrollo de las telecomunicaciones, surge otra forma de administración: presencial, cuando se aplica personalmente al sujeto o a distancia, cuando se utiliza algún medio que teletransmisión de la prueba.)

***Según el tiempo de aplicación:** De tiempo (con tiempo limitado estrictamente) y de potencia (en los cuales la mayoría de las personas terminan la prueba en el tiempo especificado) Aplicable sobre todo a pruebas cognoscitivas.

***Según el medio de administración o el material empleado:**

a) Pruebas de papel y lapiz. El sujeto contesta por escrito, con palabras, marcas, dibujos, etc. y los reactivos se presentan por escrito.

b) Pruebas orales. En las cuales la palabra hablada es el medio de preguntar y de contestar.

c) Por terminal. Una variación de las anteriores es aquella en la cual la prueba es administrada en la pantalla de una computadora y el sujeto contesta utilizando el teclado de la misma. Esto permite eventualmente la aplicación a distancia.

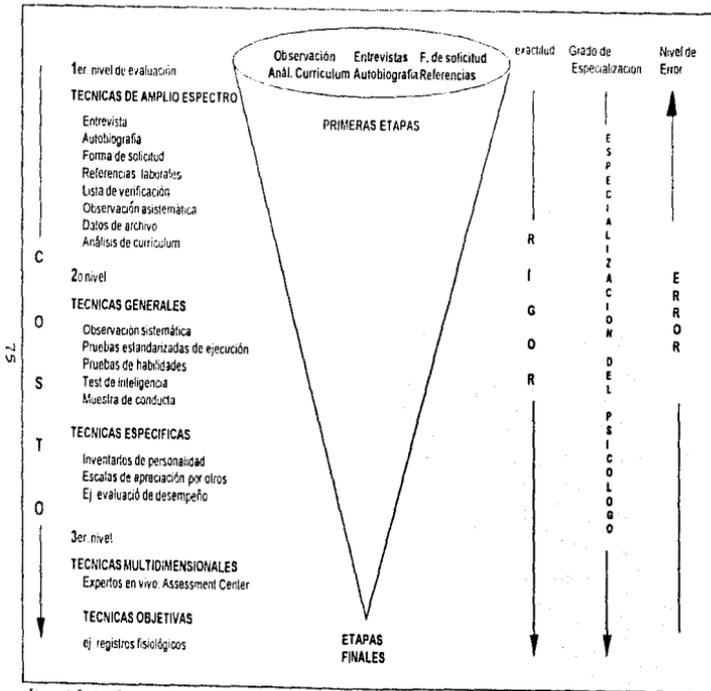
d) De ejecución. En la cual la persona tiene que operar objetos o equipo.

e) Simulación. En las cuales la persona es colocada en una situación simulada de trabajo. (ej. un juego de negocios, un role-playing o una charola de entrada.)

f) Situacionales o muestra de trabajo. en la cual la persona es colocada en una situación real de trabajo. Por ejemplo, el realizar un trabajo en el turno.

Para finalizar con este punto y en un intento de sintetizar la información, se ha preparado un esquema de aplicación de los instrumentos de evaluación y su relación con las etapas de evaluación, el rigor metodológico, el costo, la posibilidad de error y el nivel de especialización del psicólogo el cual se presenta en la siguiente página.

PROCEDIMIENTOS DE EVALUACION PSICOLOGICA
 SEGUN LA ETAPA, EL COSTO, EL RIGOR Y EL NIVEL DE ESPECIALIZACION.
 Adaptado para evaluación psicológica laboral



Esquema de Ferrando Batetxea (1982) adaptado por Marcela Fuentes (1997)

En la figura anterior, se presenta un esquema de la interrelación entre las diferentes fases del proceso de evaluación y las técnicas de evaluación psicológica con variables como el grado de especialización del psicólogo, el costo, el rigor metodológico y el nivel de error de medida.

Este esquema se preparó a partir de uno de Fernández-Ballesteros, (1992, vol 1, p 85) en que plantea una idea de "embudo" para ir ganando en especificidad, rigor y exactitud en los instrumentos conforme se avanza en el proceso de evaluación con la intención de contrastar por diferentes métodos y distintas fuentes de datos para aumentar la validez.

En una primera fase, el evaluador utiliza técnicas de amplio espectro y bajo costo, como la forma de solicitud, la entrevista, la observación asistemática, etc, pero también de alto error probable. La especialización del psicólogo requerida para esta fase es comparativamente menor que la de etapas posteriores. Aquí se generan las hipótesis iniciales relativas al caso. En el caso de selección de personal, esta etapa corresponde a la de preselección de la base del reclutamiento.

En fases posteriores, el psicólogo planea la aplicación de técnicas más costosas pero con mayor especificidad, rigor, confiabilidad, validez y menor error probable, como la observación sistemática por allegados o por especialistas, las pruebas estandarizadas, la evaluación de la personalidad, los simulacros assessment-center, etc.

En las fases finales, se constatan las hipótesis formuladas, se emiten los resultados y se toman las decisiones.

En este esquema, no se tomó en cuenta otra condición importante, a veces determinante en un contexto laboral, que es el tiempo de entrega de resultados que en ocasiones obliga a utilizar técnicas de mayor error o menor rigor.

2.3.1 LOS INVENTARIOS DE PERSONALIDAD

Los autoinformes o autodescripciones.

El autoinforme es el más antiguo procedimiento para obtener información de las personas. Implica un mensaje verbal que el sujeto emite sobre cualquier tipo de manifestación propia; se ubica en el campo de la evaluación psicológica al perseguir un objetivo definido y al usarse con un método particular. La suposición en que se basa este método es sencilla, según Brown, (1980): el individuo está en la mejor posición para observar, describir y señalar su propia conducta y su autoinforme puede abarcar la gama más amplia posible de observaciones.

Según Fernández Ballesteros(1992), el autoinforme adopta diferentes modalidades que varían en rigor metodológico y científico: la entrevista, los cuestionarios, inventarios o escalas, los autorregistros y la técnica de pensamiento en voz alta. Las siguientes son características de los autoinformes en general, mencionados por la autora

- a) La conducta autoinformada y el propio mensaje verbal son considerados como indicadores de un atributo subyacente o disposición psicológica presente en todos los sujetos en alguna cantidad o medida.
- b) Los elementos elegidos para construir este tipo de autoinforme se seleccionan a través de estrategias racionales, empíricas o factoriales y esas estrategias explican las relaciones entre los reactivos y las dimensiones medidas
- c) El contenido del autoinforme no está necesariamente relacionado con la característica que evalúa. Esto se refiere a que el contenido (verbal) del reactivo puede no tener una relación aparente con el rasgo o factor que evalúa convirtiéndose en un indicador indirecto de tal rasgo.
- d) Independientemente de la estrategia usada en su construcción, esta opera sobre la base de un diseño intrasujetos que permite la comparación con las respuestas de otros sujetos que forman el grupo normativo.
- e) Estos autoinformes parten de la hipótesis de la generalidad de la conducta y sus elementos están redactados asimismo en forma general, con el fin de identificar como el sujeto dice que se comporta habitualmente sin considerar la especificidad situacional de la respuesta.

Bellak y Hersen (1977) citados por Fernández-Ballesteros (1992) señalan que un sujeto puede informar verbalmente acerca de sus conductas motoras, de sus respuestas fisiológicas, acerca de ciertas cogniciones y acerca de sus experiencias subjetivas en relación con determinados actos o comportamientos, atribuciones, expectativas, etc. Es posible ubicar esta clasificación en un continuo que va de mayor a menor precisión en la información externada.

Los autoinformes pueden referirse a múltiples manifestaciones tanto externas como internas. Las primeras son susceptibles de comprobación; las segundas, no. Prácticamente no se cuentan con pruebas independientes que sirvan de criterio a la hora de analizar, en sus contenidos, el mundo cognitivo del individuo; para evaluar este, los autoinformes pueden ser considerados el método prioritario y directo por excelencia. De acuerdo con Leberman (1979, citado por Fernández-Ballesteros 1992) los autoinformes están siendo rehabilitados por ser la única vía de acceso a determinadas manifestaciones encubiertas del individuo.

Modelo de procesamiento de la información aplicable a autoinformes verbales.

Simon, en diversas publicaciones (Newell y Simon, 1972, Simon, 1979, Ericsson y Simon, 1980, 1984, en Fernández Ballesteros, 1992) ha planteado un modelo sobre el procesamiento de la información aplicable a los autoinformes verbales en general pero que es especialmente útil para comprender el proceso seguido al contestar inventarios de personalidad.

Se considera que cualquier información es almacenada en tres diferentes memorias: almacén sensorial de muy corta duración, memoria a corto plazo de capacidad limitada, MCP, y memoria a largo plazo, de ilimitada capacidad y almacenamiento, MLP.

La información a la cual se atiende se mantiene en la MCP y es directamente accesible para producir informes verbales entre otras operaciones. Para tener acceso a la información almacenada en la MLP, ésta ha de ser recuperada y transferida a la MCP antes de producirse la información verbal sobre la misma. Debido a la limitada capacidad de la MCP, esta sólo puede informar directamente sobre eventos que se están produciendo en el momento. Si se desea informar sobre eventos pasados almacenados en la MLP, antes se deberá recuperar esa información. La exactitud del autoinforme es menor en el segundo caso y se reduce aun más cuando se le solicita a la persona que informe acerca de elaboraciones o inferencias sobre la información almacenada, como son valores, atribuciones, expectativas, intenciones, etc.

En los casos en que se interroga a los sujetos sobre este tipo de información, registrada, ya sea en la MCP o en la MLP, ésta ha de ser elaborada por la persona y de ella inferidos conceptos. Sin perder de vista que estos datos sean relevantes, estos autoinformes son menos válidos, fiables y exactos como lo refieren Niesbett y Wilson, 1977, Ericsson y Simon, 1980, Hollon y Bennis, 1981, en Fernández Ballesteros, 1992.

Otra variable a considerar en el proceso de información de los autoinformes, agrega Fernández-Ballesteros (1992), es el momento en el que se produjo el hecho del que se informa. Según el lapso entre la ocurrencia del evento y el autoinforme del mismo, los reactivos de un inventario demandan información retrospectiva, concurrente o futura. Los reactivos retrospectivos se refieren a sucesos pasados pero pueden referirse a muy distintos tiempos pasados, a un pasado reciente o a un pasado remoto. El tiempo transcurrido entre la ocurrencia del evento y su informe requiere analizarse a la luz de los procesos de la memoria como la selección y el olvido.

La evidencia, según esta autora, sostiene que los autoinformes retrospectivos del pasado próximo son más fidedignos que los del pasado remoto. Los autoinformes específicos sobre conductas son más fieles que aquéllos que informan sobre inferencias que se pide al sujeto que realiza actualmente sobre eventos que ocurrieron en el pasado.

Resumiendo, en la medida que los reactivos demanden al sujeto información pasada o elaboraciones, inferencias, etc. de hechos pasados, presentes o futuros, el proceso que sigue la información es más largo y está más sujeto a mecanismos selectivos, olvido, etc. Como consecuencia, esta información será menos fiable y exacta.

Los inventarios como una técnica de autoinforme o autodescripción.

Los inventarios son instrumentos eminentemente psicométricos en el sentido que proporcionan calificaciones cuantitativas, que permiten la comparación entre sujetos. Se construyen a la luz de un enfoque de rasgos con una metodología racional, empírica o factorial y se utilizan para la evaluación de la personalidad, los intereses, las motivaciones, etc. Deben llenar los requisitos o propiedades de rigor mencionados anteriormente como son la confiabilidad, la validez, la estandarización, la normalización y la objetividad.

Son instrumentos verbales de autodescripción ya que obtienen información verbal a partir de un estímulo igualmente verbal; funcionan sobre la base de comunicación de la palabra escrita para mantener la estandarización de los estímulos y las respuestas. El manejo posterior de estas respuestas, uniformes en su formato para todos los sujetos, permite gran diversidad de análisis estadísticos, con la ventaja de que una extensión o revisión de éstos, amplía su vigencia en el tiempo, en países, culturas e idiomas diferentes de aquéllos de su origen. Este tipo de instrumentos es especialmente útil para evaluar grupos numerosos de personas.

Mischel (1988), menciona que una característica importante de los inventarios es el grado de estructura de la respuesta, ya que el sujeto tiene que responder con una selección de un número limitado de opciones dadas. Esta restricción en la respuesta marca una distinción sustancial con las técnicas proyectivas que permiten la libertad en las respuestas, como se mencionó en el tema de técnicas proyectivas. Esta estructuración es precisamente lo que permite manejar las respuestas de dos o más sujetos y aplicar una amplia variedad de tratamientos estadísticos.

La definición de Fernández-Ballesteros (1992) enfatiza el proceso de autodescripción: los tests (inventarios) de personalidad son autoinformes tipificados contruidos por medio de procedimientos psicométricos que permiten obtener una puntuación diferencial, es decir, la posición relativa de un sujeto en una determinada variable intrapsíquica tras la comparación de sus respuestas con las de un grupo normativo.

ESTA TESTS NO DEBE
SALIR DE LA BIBLIOTECA

En palabras de Kline (1993) los cuestionarios de personalidad consisten en un conjunto de reactivos, usualmente preguntas o afirmaciones acerca de sentimientos o conductas a los que el sujeto tiene que contestar con respuestas cerradas o indicando su acuerdo o desacuerdo con la afirmación. Agrega que son instrumentos de evaluación de la personalidad preferidos por los psicómetras debido a la relativa facilidad para construir tests confiables y establecer las normas.

Kline (1985) considera que los inventarios de personalidad factoriales son sistemas poderosos para la medida en los estudios de la personalidad y que sus problemas y dificultades pueden superarse. Agrega que los inventarios de personalidad son razonablemente discriminativos, por su longitud y consiguiente varianza y superan a la mayoría de los métodos de evaluación de la personalidad.

La preocupación por resolver problemas prácticos aplicados y el interés por una mejor comprensión teórica del ser humano, llevaron al desarrollo de los inventarios de personalidad. Goldberg (1970, citado por Lanyon y Goodstein, 1977) distingue entre dos clases de inventarios basados en las razones de su desarrollo: El primer grupo de inventarios fue elaborado en respuesta a la necesidad de tratar problemas específicos y el segundo grupo se basó en concepciones de la estructura de las diferencias individuales y se consideró como surgido de conceptos teóricos acerca de la naturaleza de la personalidad. (p 15 y 22)

Una clasificación de los inventarios de personalidad según su construcción y el tratamiento posterior de sus respuestas.

Kelly (1967 en Fernández-Ballesteros, 1992) presenta tres estrategias de construcción de inventarios: Racionales, empíricas y factoriales. En los primeros, se reúnen elementos que a juicio del investigador o según una teoría, tengan relación con el fenómeno. Aiken (1996) los llama inventarios con validez de contenido y menciona como inventarios representativos al Inventario de Preferencias Personales de Edwards y el Indicador de Tipos de Myers-Briggs. Esta técnica es la más sencilla y con mayor margen de error pero a un inventario construido así se le pueden adicionar estudios psicométricos que mejoren sustancialmente su eficiencia.

Los inventarios construidos con una estrategia empírica son aquellos cuyos elementos identifican diferencialmente entre dos grupos clasificados por otros criterios. Se eligen los ítems con mayor poder de discriminación. Ejemplos de inventarios importantes construidos empíricamente son el MMPI y el MMPI-2. Otro es el CPI, Inventario Psicológico de California, mencionados por Aiken, (1996)

La estrategia factorial para la construcción de inventarios parte de un análisis factorial. Este es un procedimiento matemático para analizar una matriz de correlaciones entre las respuestas a los reactivos a fin de determinar qué factores (constructos o dimensiones) son suficientes para explicar dichas correlaciones. (Aiken, 1996) Este método puede aplicarse a inventarios contruidos con otras técnicas.

Estas tres estrategias no son mutuamente excluyentes y con frecuencia se combinan. Aiken (1996) agrega que, en el esfuerzo por el desarrollo de nuevos y mejores instrumentos, se ha optado por una combinación de técnicas como las seguidas por D.N. Jackson al elaborar la Forma de Investigación de la Personalidad, PRF y el Inventario de Personalidad de Jackson, JPI.

DESCRIPCIONES DE ALGUNOS INVENTARIOS.

A continuación se describirán algunos inventarios utilizados en evaluación de personal, además del MMPI, MMPI-2 y el Cuestionario 16 PF que ya se describieron anteriormente.

1. Cuestionario 16 PF-5. Es la 5a. revisión del Cuestionario 16 PF (Catell, R.B, Catell A.K.S. y Catell, H.E.P.) publicado en 1993 en EE.UU. Evalúa 16 rasgos primarios y cinco factores secundarios: Extraversión, Ansiedad, Dureza, Independencia, Auto-control. Cuenta con escalas de validez ya integradas. En 1995 se publicó la traducción y adaptación a España con normas para adolescentes y adultos, hombres y mujeres. Se cuidó la redacción de los elementos y la presentación del cuadernillo de aplicación. Esta versión no puede ser calificada manualmente sino por computadora en un intento por proteger la prueba. (Seisdedos, 1995) Aún no ha sido adaptada a México

2. Perfil e Inventario de la personalidad Gordon, PPG e IPG. (2a. Ed.) Dos partes que en conjunto miden ocho rasgos: Ascendencia, Responsabilidad, Estabilidad emocional, Sociabilidad, Cautela, Originalidad, Relaciones interpersonales y Vigor. Asimismo se obtiene un factor de Autoestima. Tiempo de aplicación aprox. 30' Calificación manual. Distribuido en México. Baremos para ejecutivos, estudiantes, vendedores. (Catálogo 1997 de El Manual Moderno) Construido con metodología factorial, la estabilidad temporal oscila de .84 a 87 según factores. (Fernandez-Ballesteros, 1992)

3. Cuestionario de la Personalidad Ocupacional OPQ. Este es un inventario construido en Gran Bretaña en un esfuerzo conjunto de una firma, SHL y un grupo de 53 corporaciones importantes. Desde su concepción inicial, fue enfocado a la personalidad ocupacional Publicado en 1984, actualmente es usado en todo el mundo, especialmente en EEUU y Gran Bretaña. Está disponible en 12 idiomas Este instrumento esta completamente computarizado, en su aplicación, interpretación y generación de informes narrativos lo que restringe su uso por las empresas o consultores autorizados que adquieren el sistema y reciben capacitación lo que permite proteger el instrumento y evitar usos indebidos.

Mide 31 rasgos agrupados en tres categorías: relaciones interpersonales, estilos de pensamiento y sentimientos y emociones Con base en estas dimensiones, se evalúan aspectos importantes para el trabajo así como estilos de liderazgo y de subordinación, comportamiento en equipo y estilos de influencia y venta

El Cuestionario de la Personalidad Ocupacional OPQ está adaptada a México y cuenta con normas para directivos y profesionistas. Una ventaja es la generación de informes computarizados que permite evaluar grupos numerosos y entregar resultados en poco tiempo. Se complementa con dos pruebas de habilidades: verbal y numérica, igualmente estandarizadas en México. Asimismo, cuenta con un sistema perfilador de puestos que permite obtener una estimación de la congruencia persona-puesto en las mismas dimensiones de la prueba. (Información proporcionada por SHL de México.) Furham (1992) refiere confiabilidad adecuada.

4. Inventario Psicológico de California CPI (Gough, 1957, 1975, 1987) Construida con criterio empírico y racional, es como un MMPI para población normal, del que fue derivado y es similar en su formato. Consta de 480 ítems en 18 escalas y tres de validez. (Fernández-Ballesteros, 1992) Proporciona información en cuatro categorías. a) Actitud, ascendencia, seguridad de sí mismo, adecuación. b) Socialización, madurez, responsabilidad, estructuración personal de valores. c) Potencial para el logro y eficacia intelectual. d) Modos intelectuales e intereses (Catálogo de El Manual Moderno) El manual editado en México no informa si las normas corresponden a nuestro país.

Criticas y limitaciones de los inventarios de personalidad

Kline (1993) enumera una serie de problemas o desventajas de los inventarios de personalidad:

1. El número de variables a medir es uno de los problemas que tienen que resolver los constructores de inventarios. Los psicólogos de los rasgos y especialmente los factorialistas no se han puesto de acuerdo en qué tantas variables son representativas e importantes para concebir la personalidad.
2. Los ítems o reactivos de los inventarios de personalidad resultan inevitablemente simplistas y es imposible capturar la rica sutileza de los sentimientos humanos en las breves afirmaciones o preguntas. Heim (1975 en Kline,1993) sostiene que se obtiene una pobre cooperación de parte de los sujetos al enfrentarlos a preguntas que tienen que ser contestadas con "sí o no" o al manifestar su acuerdo o desacuerdo con una afirmación simple.
3. El problema de los grupos y tendencias de respuesta o, más general, el falseamiento de las respuestas que es el tema del siguiente capítulo.
4. Otros problemas que tienen que enfrentar los constructores de inventarios de personalidad:
 - a) El muestreo adecuado del universo de ítems, esto es, que los reactivos de un inventario sean representativos de un universo hipotético e infinito de ítems.
 - b) La representatividad de la muestra de sujetos, en relación con la estandarización del test.
 - c) La validez de los tests de personalidad y la dificultad de establecer criterios externos claros para calcular la validez concurrente o predictiva: su validación es cuestionable. Existen tests de personalidad potentes en sus propiedades psicométricas pero con una evidencia reducida de su validez.

LAS VARIABLES LINGÜÍSTICAS

En otro orden de ideas, la influencia de variables lingüísticas es un aspecto poco estudiado en los inventarios de personalidad, que son instrumentos eminentemente verbales. Uno de los grandes aciertos del cuestionario 16 PF de Catell es su derivación de la esfera del lenguaje lo cual de alguna forma garantiza el abarcar la esfera total de la personalidad pero por esta misma razón, existe una especial necesidad de estudios lingüísticos de tipo semántico para verificar que el significado del reactivo se mantenga al cambiar los escenarios de aplicación, por ejemplo, al traducirlo a otros idiomas, al usarlo en diferente época con otra generación de sujetos, etc. Esto también se aplica a las respuestas, como lo plantean Lanyon y Goodstein, 1977, en relación a la existencia de diferencias subculturales en el uso de palabras evaluativas tales como "frecuentemente" y "muy", que pueden afectar las respuestas de la prueba pues cada persona puede entenderlo de diferente manera.

Otro riesgo del lenguaje en la evaluación por inventarios factoriales es el creer que las escalas de diferentes inventarios que tienen un mismo nombre o un nombre similar, necesariamente miden lo mismo. Recuerdese que los investigadores nombran un factor después de identificarlo. Aiken (1996) recomienda que ni siquiera debe considerarse así en los casos en que existan correlaciones altas entre tales escalas homónimas, porque estas correlaciones pueden ser ilusorias. Nunally (1987) va más allá al mencionar que los inventarios están plagados de severos problemas semánticos, que se evidencian tanto en la comunicación del significado de los reactivos a los sujetos como en la comunicación de los resultados de la evaluación.

Otro problema lingüístico se produce artificialmente al intentar reducir la tendencia exagerada a estar de acuerdo con los reactivos o condescendencia. La estrategia de cambiar la direccionalidad de los elementos y redactarlos en una forma negativa aumenta la dificultad para comprender el significado del ítem y el significado de una respuesta cuando se junta un doble negativo. Una innovación en el 16 PF-5 es que adiciona un refraseo del reactivo negativo a la respuesta indicando que significa la respuesta verdadero o falso, lo que facilita la comprensión del contenido. (Catell et al, 1993, Seisdedos, 1995)

El punto anterior nos lleva a considerar la cuestión de la aplicabilidad de los inventarios a otras culturas. De acuerdo con Price-Williams (1961 en Kline, 1985) es sumamente difícil desarrollar cualquier test genuinamente transcultural. El contenido de los reactivos usualmente tiene bastante influencia de la cultura en que se origina. Kline (1985) pone como ejemplo uno de los elementos del AI3Q (Kline, 1971) que trata con salir a cenar fuera de casa. Esta prueba se aplicó en la India, África y Gran Bretaña, lugares en que la gente sale a cenar, pero en circunstancias tan diferentes entre sí que cualquier comparación transcultural resultaría un absurdo.

Otra cuestión importante, esta de orden teórico, (Seisdedos 1985) es el planteado por pretender perfilar todo un mapa de configuraciones más o menos estables, que son los rasgos fundamentales de la personalidad a partir de unidades de comportamiento, las descritas en los reactivos de los inventarios de personalidad. (p 25) Este tema se tocó al revisar la teoría de los rasgos y las teorías factoriales de personalidad en el primer capítulo

Los inventarios en la evaluación de personal.

De los criterios para elegir instrumentos de evaluación que refieren Lanyon y Goodstein (1977) que son el factor tiempo y la economía, al referirlos a los inventarios se convierten en una ventaja. El primer punto se refiere tanto al tiempo de aplicación, calificación e interpretación, al tiempo de entrega de resultados, al tiempo que el evaluador requiere de entrenamiento e incluso, el tiempo que dedica el "cliente" dentro de la organización para leer y comprender el informe. En este aspecto, los inventarios tienen amplia ventaja sobre otros métodos de evaluación excepto la entrevista, pero tienen a su favor el ser instrumentos objetivos, estandarizados, etc. Es factible realizar aplicaciones grupales y obtener resultados rápidamente. Este tipo de tests son además fácilmente adaptables a uso computarizado.

Los inventarios son asimismo instrumentos de costo relativamente bajo, por diversas razones. El precio de los inventario rara vez es exagerado y generalmente los cuadernillos de aplicación son reusables. Debido a que es reducido el tiempo del evaluador en la aplicación, etc, el costo por este concepto es también bajo, comparado con otras técnicas como las proyectivas o los simulacros assessment-center. Asimismo, el costo del entrenamiento requerido para el manejo de inventarios es comparativamente menor.

Las características psicométricas constituyen una ventaja por si mismas. Los resultados permiten comparaciones de muy amplia cobertura: entre individuos, intraindividuos, entre grupos, entre perfiles de puesto-persona, etc. Sin embargo, no es conveniente caer en el extremo de tomar los resultados de un inventario como única base para una decisión trascendente como una contratación o promoción de personal. Los estudios sobre las diferencias entre perfiles de grupos profesionales, evaluados con el 16 PF, no son lo suficientemente discriminativos para basarse solamente en este test (Kline, 1985) y no es posible pretender que un sujeto, con un conjunto dado de puntuaciones pueda, automáticamente, ser seleccionado para un puesto específico.

Los resultados son mejores si los inventarios se combinan con otras técnicas, por ejemplo, con fines de selección de personal, puede complementarse con un formato de solicitud ponderado, con entrevista estructurada y pruebas de capacidades.

Uno de los procedimientos de validación de inventarios en el campo laboral consiste en correlacionar las puntuaciones de los rasgos con alguna medida de éxito en el empleo y determinar el valor de los inventarios como predictores laborales. (Kline, 1985) Este tipo de validación requiere de criterios muy bien definidos y cuantificables con los cuales asociar los puntajes de los inventarios.

Por la información anterior, es posible concluir que los inventarios, convenientemente manejados, son herramientas útiles y versátiles en la evaluación de personal, que permiten además la investigación y la revisión constante pero que tiene la limitante de su susceptibilidad a la distorsión al utilizarse en selección de personal. (Catell y Kline, 1982) Esta distorsión de las respuestas es el tema del siguiente capítulo.

CAPITULO 3.

EL FALSEAMIENTO DE LAS RESPUESTAS EN LOS INVENTARIOS DE PERSONALIDAD.

3.1 EL ERROR DE MEDIDA

Todo puntaje obtenido con un instrumento psicológico se considera como la suma de un puntaje verdadero más un error de medición; este error puede ser un sesgo sistemático o un error aleatorio. La prevención sistemática no es tan importante en la mayoría de las medidas psicológicas como lo son los errores aleatorios. Los errores sistemáticos sólo indican la puntuación promedio de los sujetos pero los errores aleatorios limitan la eficiencia con que se pueden descubrir los principios y las leyes. Aunque los errores aleatorios no se puedan eliminar por completo, los esfuerzos deben estar encaminados a reducirlos al mínimo. Nunnally (1987)

De mi experiencia personal en el campo de la evaluación de personal, he identificado diferentes fuentes probables de error que en seguida presento en una clasificación que no pretende ser exhaustiva.

Las principales fuentes de error en la evaluación mediante pruebas psicológicas son: Las debidas a la estructura de la prueba, las originadas en el proceso de evaluación y calificación, las debidas al evaluando y las debidas al aplicador

Error debido a la estructura de la prueba misma.

En el capítulo anterior se menciona la importancia de la validez, la confiabilidad, la objetividad y la estandarización como requisitos indispensables de cualquier instrumento de evaluación psicológica. La falta de cualquiera de ellos determina una fuente de error importante. A continuación se enumeran las fuentes de error más usuales.

Estandarización. La prueba debe contar con procedimientos estándares de aplicación y calificación específicamente definidos en su manual.

Asimismo, debe contar con tablas o baremos para interpretar las calificaciones y la población de estandarización debe ser la misma o muy similar a aquélla a la que pertenece el evaluado. Estas normas deben ser recientes para mantener su vigencia. Esta es una fuente de error común en México dado que no se cuentan con estandarizaciones de los instrumentos que se usan.

Traducción y adaptación. Las pruebas que se usan en México son pruebas diseñadas y elaboradas en otros países, con frecuencia construidas originalmente en otros idiomas y para idiosincracias distintas a la nuestra. Las traducciones de los tests verbales en algunos casos son realizadas en el extranjero lo cual ocasiona problemas de orden semántico, que pueden afectar de una manera determinante los resultados.

Validez y confiabilidad. Las pruebas o tests deben estar respaldados por estudios que acrediten su validez y su confiabilidad, en México o al menos en Latinoamérica. Un error importante en la utilización de pruebas con propósitos laborales es que la prueba NO CUBRA las necesidades de medición. Es necesario identificar -previamente a la evaluación- las dimensiones que se pretenden medir.

Error debido al proceso de evaluación y calificación.

Material de la prueba. En México existe práctica a todas luces no recomendable, de utilizar material no original lo cual propicia sesgos importantes en la apreciación de los reactivos de parte del examinado pues estas diferencias con frecuencia obran en su contra. Se reproduce el material y en ocasiones son fotocopias de enésima generación. Suele suceder que esta reproducción se hace dentro de la misma empresa y al poco tiempo circulan las copias sin control entre personal no autorizado. Se dan casos en que se utiliza material de aplicación incompleto o modificado del original

Las condiciones de aplicación. El lugar: la iluminación, el ruido, los distractores, los lápices, gomas, el cronometro, etc. El mobiliario, la privacidad, la cercanía entre evaluados.

Error de medida debida al examinado

Las siguientes son condiciones usuales o circunstanciales del examinando que afectan la evaluación. El evaluador debe ser perceptivo y debe decidir en que casos debe cancelar la aplicación.

- Baja capacidad para comprender la prueba. Capacidad intelectual y razonamiento verbal que limite la comprensión de la prueba.
- Condiciones físicas circunstanciales o temporales: fatiga, desvelo, problemas personales, ansiedad, depresión por desempleo, estados de confusión, hambre, gripe u otras enfermedades, presión de grupo, medicación o alcoholismo, etc.
- Limitaciones físicas de visión, oído, percepción, exceso de nerviosismo, dislexia, distalía, etc.
- Velocidad de razonamiento y ritmo de trabajo demasiado cuidadoso o demasiado rápido.
- Conocimiento previo de la prueba.
- Escolaridad inferior a la establecida por la prueba.
- Trampas del examinado. Copia o comunicación con otros evaluados, adelantos, etc.
- Inseguridad y timidez para plantear dudas acerca de las instrucciones.
- Distorsión de la prueba. Defensa, evasión, distorsión motivacional, distorsión negativa, defensas inconscientes, resistencia al aplicador, respuestas descuidadas o al azar, etc.
- Motivaciones y actitudes del examinado. Relación interpersonal establecida con el aplicador.
- Idioma. Las pruebas deben estar en el idioma materno del evaluado.

Error debido al examinador.

- Motivaciones y actitudes del examinador, simpatía o antipatía por algún evaluado.
 - Compromiso o presión para que resultados individuales o grupales sean superiores a los reales.
 - Instrucciones. Aunque las instrucciones forman parte de la estructura de la prueba, cuando éstas son orales, son responsabilidad directa del aplicador. Se debe apegar a las instrucciones especificadas en el manual para mantener condiciones de estandarización. Errores en este punto pueden ser: instrucciones confusas, inconsistentes, falta de verificación de la comprensión de estas, etc.
 - Manejo de la ansiedad propia y de los examinandos Control del grupo en las aplicaciones colectivas y evitación de trampas tanto del examinado como del aplicador que puede favorecer o perjudicar a cierto evaluado
 - Falta de capacitación en el manejo de la prueba.
 - Calificación e interpretación errónea o sesgada. Subjetividad
- En cuanto a resultados e informe, los errores pueden ser de fatales consecuencias:
- Un diagnóstico inexacto o incongruente con las necesidades organizacionales.
 - Confusión en la forma de presentar los resultados
 - Falta de apoyo a usuarios en la comprensión del informe.

3.2 EL FALSEAMIENTO O DISTORSION DE LAS RESPUESTAS.

Como se mencionó en el capítulo anterior, una de las más importantes críticas a los inventarios de personalidad es su susceptibilidad al falseamiento de las respuestas, especialmente en escenarios aplicados como la evaluación de personal (Christensen 1994) y es este el propósito del presente capítulo.

Germain (1971, en Seisdedos, 1985) plantea la interrogante de cómo separar la varianza de la personalidad de la varianza de la metodología. Esta pregunta es sustancial al intentar una evaluación -y consiguiente comprensión- de la personalidad. Megargee (1971) por su parte, señala que una distorsión de respuesta " puede definirse burdamente como un factor distinto del contenido de los elementos que influyen en la respuesta de una persona a un test" (en Fernández-Ballesteros, 1992)

Las respuestas a los procedimientos de evaluación de la personalidad están influidas por otras variables además de las características de personalidad de los respondentes. Aunque se supone generalmente que estas son las determinantes principales de sus respuestas, se sabe, sobre la base de la investigación y de la teoría, que estas respuestas son productos complejos de una serie de variables psicológicas, sociológicas, lingüísticas y de otras clases, muchas de las cuales nada tienen que ver con los propósitos para los que fue ideado el procedimiento de evaluación.

Es necesaria la comprensión de cómo estas irrelevancias sistemáticas afectan las respuestas a los instrumentos de evaluación de la personalidad, tanto para mejorar la eficiencia al utilizar los instrumentos actuales, como para desarrollar nuevos. Lanyon y Goodstein (1977)

El estudio de las distorsiones de respuesta en pruebas de rendimiento lo inicio Cronbach (1941, 1942, 1946 y 1950 en Lanyon y Goodstein, 1977) quien despues extendio este estudio a los instrumentos de personalidad tanto a los autodescriptivos como a los proyectivos como el Rorschach, el TAT y otros

Lanyon y Goodstein (1977) consideran que las distorsiones mismas de respuestas son consistencias en la conducta que a veces permiten inferencias utiles acerca de la personalidad de los respondentes que las manifiestan. El grado de utilidad estriba en la base empirica de dicha inferencia. Las respuestas que no se relacionan con las conductas evaluadas por la prueba, constituyen, por definicion, la varianza de error, y deben hacerse grandes esfuerzos para eliminarlas o reducir las substancialmente

Mischel (1988), refiere a este respecto que las preguntas y las instrucciones de las pruebas de autoinforme con frecuencia exigen que el respondente vaya mucho mas alla de la observacion directa de la conducta y que haga inferencias subjetivas sobre el significado psicologico de la conducta. Es conveniente repetir aqui lo referido por Seisdedos (1985) el señala que cuando a un sujeto se le pide que conteste al conjunto de reactivos de un cuestionario, se le esta requiriendo que interactue con cada uno de ellos. Agrega Mitchel (1988) que, aunque las preguntas están impresas y son invariables, por lo general son vagas y ambiguas; requieren que el sujeto haga una evaluacion de la conducta y luego realice una generalizacion de la misma, en vez de que los reactivos describan conductas especificas, situadas dentro de contextos particulares sobre dimensiones claras.

Ademas, prosigue Mitchel, es factible que las personas no estén dispuestas a revelarse tal como son o que no sean capaces de hacerlo, en respuesta a esos reactivos, sobre todo cuando sufren perturbación emocional o se percatan de que las respuestas emitidas pueden servir para hacer decisiones importantes sobre ellas mismas, como es el caso de una evaluacion en un contexto laboral. Todo lo anterior puede ser fuente de eventual distorsion en los inventarios

Otro punto de vista es el de Brown, (1980), quien analiza la distorsión de las respuestas en inventarios a la luz de la clasificación que divide a las pruebas en medidas de ejecución máxima y medidas de ejecución típica. En las primeras la meta es obtener un indicador del mejor desempeño posible del sujeto, se ocupan de las funciones cognoscitivas: habilidades, capacidades, conocimientos, etc y los unicos modos de falsear o distorsionar los resultados son: el negarse a responder, el elegir deliberadamente las respuestas erróneas o el adivinar.

En las segundas, lo que interesa es el comportamiento habitual y se pretende evaluar los intereses, sentimientos, actitudes, reacciones. En las pruebas de ejecución típica es mayor la probabilidad de distorsión o falseamiento y esta adopta diversas formas que serán analizadas mas adelante en las clasificación de los principales tipos de distorsión

ENFOQUES TEORICOS SOBRE EL FALSEAMIENTO.

EL SITUACIONISMO.

La mayoría de las teorías de la personalidad suponen que el comportamiento es congruente o consistente a través del tiempo y de las situaciones. Si una persona es agresiva, por ejemplo, tenderá a serlo en una amplia gama de situaciones y continuará así de un año a otro. Este comportamiento constantemente agresivo es indicador de un rasgo de personalidad subyacente de personalidad. En relación a este supuesto, algunos teóricos se han preguntado que tan consistente es la conducta; opinan que esta idea de la congruencia es ilusoria, que se ha recurrido a ella a pesar de que las pruebas empíricas de esta sean escasas y que las situaciones variables a menudo explican más sobre la personalidad que los propios rasgos. (Kelley, 1955, Mischel, 1969 en Morris, 1992) Esta perspectiva se ha llamado Situacionista y uno de sus principales exponentes es Walter Mischel quien revisó estudios sobre personalidad y encontró que la congruencia del comportamiento en diferentes situaciones en realidad era muy baja y que la gente actúa de manera distinta en diferentes situaciones. (Mischel 1968 en Morris, 1992) Concluyo que cualquier análisis de personalidad que considere únicamente las disposiciones internas o rasgos, debe ser considerado -al menos- como inadecuado (1977, citado por Morris, 1992) Considera que la aparente consistencia del comportamiento es resultado de una observación o evaluación en un limitado número de situaciones.

Mischel (1988), plantea el estudio de una interacción de rasgos y situaciones como una aproximación adecuada a los rasgos. Agrega que debe examinarse con seriedad el modo en que las cualidades de la persona y la situación influyen entre sí, es decir, "su interacción" (Bowers, 1973, Endler, 1973, Magnusson, 1980, en Mischel, 1988). Las expresiones de los rasgos de una persona dependen de su situación psicológica del momento. Refiere que en la práctica es evidente el hecho de que las calificaciones sobre cualquier rasgo se ven afectadas por todo tipo de variables irrelevantes para el rasgo que se pretende medir con la prueba, por ejemplo, Masling (1960, Mischel, 1968, en Mischel, 1988) encontro que variables como el sexo del examinador, el sexo del sujeto o la amigabilidad, influyen de alguna manera en las calificaciones. De parte del sujeto también existen peculiaridades que pueden afectar los resultados. Las pruebas sólo generan una muestra de conducta bajo determinadas condiciones y la conducta observada siempre ocurre dentro de un contexto específico, es decir, dentro de una situación psicológica

Morris (1992) menciona que Mischel posteriormente suavizó su posición y entonces considero que al menos algunos comportamientos son relativamente consistentes a través del tiempo y de diferentes situaciones. Las investigaciones señalan que algunas personas son mas consistentes que otras.

Bandura (1977 en Morris, 1992) suma al binomio rasgos-situación, otro factor: la retroalimentación que se reciba del comportamiento real de la persona en una situación dada. Las personas evalúan una situación conforme ciertas expectativas internas y esta evaluación produce cierto tipo de conducta. La retroalimentación ambiental que recibe la persona de la conducta modela, a su vez, las expectativas conductuales de la persona para el futuro. Al aplicar estos conceptos a una evaluación de rasgos mediante inventarios, debe considerarse que el respondiente está contestando en función de las características más o menos consistentes de su personalidad, de la situación y de las expectativas que tenga de esa evaluación en especial. Mischel (1986 en Aiken, 1996) se suma a ese planteamiento de aprendizaje social y lo prefiere en lugar de analizar la personalidad simplemente por un conjunto de rasgos o factores.

En relación a este punto, Fernández Ballesteros (1992) agrega que la posición situacionista enfatizó la especificidad del comportamiento a diferencia del concepto de la generalidad de la conducta, es decir, que este es dependiente de los estímulos o situaciones presentes. Así, por ejemplo, la ansiedad hace referencia a un complejo patrón de respuestas cognitivas, psicofisiológicas y motoras. Este concepto puede ser visto como una disposición interna que provoca que el sujeto que la ostenta se comporte generalmente de una manera ansiosa; pero también puede ser entendido como un estado en el sentido de que la persona pueda presentar ansiedad en determinadas situaciones. Spielberger (1966, 1991, citado por Fernández-Ballesteros, 1992), ha enfatizado la importancia de la distinción entre rasgo-estado y con base en estos conceptos construyó un test con dos formas que identifica tanto el rasgo como el estado de ansiedad.

La teoría de los rasgos debe considerar, en sus métodos de evaluación, la influencia de la situación del evaluado, que pueden propiciar variables diferentes a aquellas consideradas en su modelo. En este caso, los inventarios estarán expuestos a ser influidos dependiendo de la situación en que sean usados. Mischel (1968 en Aiken, 1996) postula que al parecer las variables afectivas están más influenciadas por los factores de situación que las variables cognoscitivas; concluyó en ese entonces que el comportamiento personal y social depende en gran medida en la situación específica y que ésta es de suma importancia para predecir el comportamiento.

Por otro lado, investigadores como Bem y Allen, 1974; Block, 1977; Underwood y Moore, 1981; Chaplin y Goldeberg, 1984, (citados por Aiken, 1996) llegaron a la conclusión que la consistencia de los rasgos en todas las situaciones es en sí una variable de diferencia individual. Por su parte, Funder y Colvin, (1991 en Aiken 1996), consideran que algunos comportamientos son más consistentes que otros y que ciertas conductas son específicas de la situación en tanto que aquellas que no requieren de un estímulo específico se presentan en gran variedad de situaciones y reflejan mejor las variables amplias de la personalidad.

Mc Reynolds, (1979, citado por Aiken, 1996) indica que los esfuerzos futuros por desarrollar medidas efectivas de la personalidad deberían empezar por proporcionar un modelo conceptual de la forma en que los rasgos y las situaciones interactúan y después construir medidas de ambos grupos de variables para hacer una verdadera "evaluación de interacción"

LA TEORIA DE LA PERSPECTIVA DEL RASGO.

Catell (1973 en Catell y Kline, 1982) aborda las distorsiones motivadoras mediante una teoría mucho más general. El plantea que la deseabilidad social nunca se presenta como factor de primero o segundo orden y que se necesita un modelo distinto para identificarla, ya que la puntuación de las escalas de la personalidad varía cuando los sujetos intentan mostrarse de un modo más atractivo. Esto lo explica mediante la teoría de la Perspectiva del Rasgo, que sostiene que toda distorsión de percepción y evaluación en los datos Q y L tiene que entenderse como producto de la personalidad del observador y su situación. Así también la percepción distorsionada, ocasionada por la deseabilidad social, puede entenderse como cualquier otro comportamiento, en términos de la ecuación de especificación (ver cap 1).

La perspectiva del rasgo explica que los factores de personalidad retienen su identidad y su verdadera forma en todas las situaciones y que la distorsión en cuanto a deseabilidad social actúa sobre ellos como entidad unitaria, aumentando o disminuyendo sus niveles de manifestación. Esto implica que las respuestas del sujeto están mediatizadas por su propia personalidad, el modo como se autopercibe y la situación motivante o neutra en que realiza el examen. Seisdedos (1985) La teoría de la perspectiva del rasgo insiste en enfatizar que cualquier método de evaluación psicológica que se base en los juicios de un observador está fuertemente influenciado por las características de dicho observador, no importa si es una persona evaluando su propia conducta mediante cuestionarios de autoinforme o si es un profesional entrenado evaluando la conducta de otros (Catell, 1968. Krug y Catell, 1971 en Krug, 1978)

Los autores partidarios de la "explicación del rasgo" de esta distorsión de respuesta argumentan que la capacidad de simulación es una importante variable de personalidad. Independientemente si se considera al falseamiento como un rasgo en sí mismo o como un factor de error, el usuario de inventarios debe tener presente el probable falseamiento y solicitar la cooperación del respondiente al dar las instrucciones con el propósito de reducir la distorsión al mínimo. Asimismo, el psicólogo debe contar con elementos para identificar el falseamiento y considerarlo en la interpretación de los resultados. (Fernández-Ballesteros, 1992)

Los inventarios de personalidad son efectivamente sensibles a la distorsión o falseamiento de las respuestas de parte del evaluado pero tienen la ventaja de que, por su misma naturaleza, esta distorsión puede ser detectada y evaluada en su intensidad y en su direccionalidad. (Krug, 1978)

EL FALSEAMIENTO DE LAS RESPUESTAS EN OTROS INSTRUMENTOS.

Como se mencionó en el capítulo anterior, los inventarios de personalidad son especialmente sensibles a la distorsión de las respuestas, lo cual sucede en menor medida en otros métodos de evaluación. Los tests objetivos o datos T, desarrollados por Catell y sus colaboradores, cuyo verdadero propósito queda oculto al sujeto, están aparentemente libres de distorsión de respuesta. La resistencia a la falsificación, la resistencia a las tendencias de respuesta y la resistencia a la autoilusión, constituyen ventajas manifestadas por Kline (1985) de este tipo de tests y son especialmente importantes en la evaluación de candidatos con fines de selección de personal. Sin embargo, estos tests existen en limitado número y aun no se han completado los estudios de validación como para usarlos regularmente para tal fin.

Lanyon y Goodstein (1977) refieren que, contrariamente a la creencia popular, las técnicas proyectivas son también muy vulnerables a los esfuerzos de los respondientes por sesgar sus respuestas, si bien, parece más difícil dar deliberadamente una buena impresión que una mala.

A este respecto, Bellak (1990) refiere que el falseamiento de las pruebas proyectivas puede ser comprendido y deben conocerse sus alcances al analizar la producción del sujeto desde el punto de vista del yo, en la medida en que se pueden introducir actitudes conscientes y un control consciente del yo, el sujeto puede llegar a alterar el protocolo en un grado mínimo, en una u otra dirección, porque estas variables forman parte de un todo. Sin embargo, Bellak considera que no es posible que esta alteración sea sistemática.

Las técnicas proyectivas no están tan exentas de distorsión como se ha manejado. En un antiguo experimento de Davids (1955 en Megargee, 1971) se encontró que es factible falsear las respuestas ante instrumentos proyectivos. Davids trabajó con dos grupos de estudiantes a los que aplicó dos técnicas proyectivas y obtuvo además otras medidas de mala adaptación. El primer grupo fue reclutado con el popular argumento de contribución a la ciencia y se encontraron correlaciones moderadamente altas entre las dos fuentes de datos. Al otro grupo de estudiantes se le hizo creer que el propósito del estudio era seleccionar un individuo de excepcional estabilidad emocional y madurez para un trabajo muy bien pagado y en este caso, las correlaciones disminuyeron considerablemente. Megargee (1971) agrega que es difícil explicar estos cambios como ajenos a la voluntad del sujeto y que el sujeto tiene un amplio conocimiento de lo que "proyecta" en una técnica proyectiva.

La habilidad para falsear un test de personalidad en parte depende del conocimiento del evaluado del papel que va a ser simulado. (Kroger & Turnbull, 1975, 1970, Foulds & Warehime, 1971, Lawton, 1963 en Albert, S. et al, 1980). Albert y sus colaboradores (1980) aplicaron Rorschach a dos grupos de estudiantes y uno de psicóticos diagnosticados como esquizofrénicos paranoides.

Los estudiantes pasaron un MMPI como filtro pero no se les había aplicado el Rorschach antes. A los estudiantes se les pidió que simularan en sus respuestas a un esquizofrénico paranoide y a la mitad de ellos se les instruyó acerca de la esquizofrenia paranoide. Los expertos en Rorschach que fungieron como jueces no fueron capaces de identificar los protocolos falseados de individuos normales, tanto de los que sabían muy poco acerca de las psicosis como de los que fueron informados pero estos últimos recibieron más designaciones de psicosis que los otros dos grupos, incluidos los psicóticos reales. Más aún, los jueces detectaron muy poca evidencia de falseamiento aun cuando éste se les presentó como una opción.

3.2.1 TIPOS DE DISTORSION DE RESPUESTAS EN LOS INVENTARIOS DE PERSONALIDAD

La siguiente es la clasificación presentada por Fernández-Ballesteros(1992) que integra las de varios autores.

1	TENDENCIAS O ESTILOS DE RESPUESTAS. Independientes del contenido de los reactivos. Se responde más al formato de la respuesta. Ej. Condescendencia, respuestas al azar.
2	GRUPOS DE RESPUESTA. Determinadas por el contenido del reactivo. Tendencia a elegir reactivos en forma congruente con una imagen social. Ej. Desestabilidad social.
3	SIMULACION O DISTORSION. Determinadas por el contenido del reactivo. Falseamiento deliberado. Ej. Distorsion motivacional.

* La mayoría de los autores llaman Agradescencia a la tendencia a estar de acuerdo pero aquí se ha preferido usar el término asentimiento o condescendencia, que se consideran traducciones más apropiadas.

ESTILOS DE RESPUESTA

Esta tendencia tiene su origen en el formato de la respuesta y es independiente del contenido del reactivo. En palabras de Nunnally (1987), son hábitos para responder las pruebas que se caracterizan por ser a) una fuente confiable de varianza en las diferencias individuales b) un producto artificial en los métodos de medición, c) por lo menos parcialmente independiente del rasgo que se pretende medir.

En el caso de respuestas dicotómicas, falso verdadero, la distorsion más frecuente es la tendencia a responder afirmativamente independientemente del contenido. Esta tendencia se denomina asentimiento o condescendencia. En otro caso, en respuestas tipo escala, la tendencia puede presentarse a contestar más bien en el centro o más bien en los extremos, independientemente del contenido de las respuestas formuladas. (Fernández-Ballesteros, 1992)

Lanyon y Goodstein, (1977) refieren que una respuesta positiva puede resultar del hecho de que un respondiente ciertamente se comporta de esa manera pero puede también significar que no está seguro de su conducta y su respuesta ha sido determinada por un estilo de respuesta de condescendencia.

Catell y Kline, (1982) refieren que la condescendencia es un grupo de respuestas comunes e insistentes, discutido por primera vez por Cronbach (1946, 1950) Esta es en sí misma un rasgo contaminante de muchas escalas del MMPI. Como ocurrió con la deseabilidad social, debe considerarse que la condescendencia influye en la medida de otros rasgos y que, como rasgo de personalidad, es interesante por sí misma.

McGe (1962 en Seisdedos, 1985) encontró tendencia a la condescendencia en los casos en que el tiempo de aplicación se reducía al mínimo. Por su parte, Norman (1963 c en Lanyon y Goodstein, 1977) identificó una dimensión de condescendencia de tal importancia que la incluyó como uno de los cinco factores básicos en su estudio factorial de la personalidad.

Lanyon y Goodstein (1977) han concluido que los estilos de respuesta es un tipo de distorsión relativamente poco importante en la mayoría de los esfuerzos por evaluar la personalidad. Rorer, (1965, en Brown, 1980) encontró que los estilos de respuestas rara vez explican más del 20% de la varianza en las calificaciones.

RESPUESTAS ATÍPICAS. Otro estilo o modo de responder a los reactivos de las pruebas de personalidad que ha atraído considerable interés es el de dar respuestas atípicas o inusitadas. Se identifican ciertas respuestas modales o comunes a los reactivos que son dadas por la población en general. Se supone que las variaciones o desviaciones de estas respuestas comunes o modales a los reactivos son indicativas de una tendencia general hacia la desviación y ha llevado a algunos autores como Berg (1955 en Lanyon y Goodstein, 1977) a estudiarla con mayor profundidad. Esta tendencia es la que origina la escala de Azar del 16PF y la escala F del MMPI.

GRUPOS DE RESPUESTA: LA DESEABILIDAD SOCIAL

Es la tendencia a responder a los elementos dependiendo de su grado de Deseabilidad Social. La deseabilidad o aceptación social, ha sido ampliamente estudiada. Edwards, (1965 citado por Fernández-Ballesteros 1992), fue el primero en demostrar que las respuestas de autoinformes clásicos como el MMPI o el CPI podían ser explicadas más en función de ser descripciones deseables o no deseables desde el punto de vista social que de saturar factores psicológicos tales como "fuerza del yo" o "ansiedad". Asimismo, la influencia de la deseabilidad social también se ha detectado en autoinformes conductuales. (Fernández-Ballesteros, Pérez Pareja y Macía, 1981 en Fernández-Ballesteros, 1992)

Lanyon y Goodstein usan el término Deseabilidad Social para referirse a la producción natural de respuestas socialmente deseables. Esta no es una distorsión deliberada sino que está en función de los valores aceptados socialmente.

Gran parte de la varianza de las mediciones de los rasgos de la personalidad derivadas de autoinventarios se puede explicar por un factor vinculado a la tendencia a decir de uno mismo cosas buenas y no malas. Edwards, 1964 resume estos conceptos denominados como aceptación social o autoconcepto expresado Nunally (1987)

Ahora bien, si se está interesado en la esfera de la personalidad de forma total, la tendencia a responder de una manera deseable desde el punto de vista social se convierte en un objeto valioso de investigación en si mismo, de tal forma que no debemos procurar eliminarlo aunque se podría restringir su influencia a un grupo determinado de elementos: una escala de deseabilidad social. Catell y Kline (1982)

Fernández-Ballesteros (1992) plantea una significativa covariación entre la deseabilidad social y las dimensiones, los rasgos, factores o repertorios de conducta que impliquen ajuste personal. Por esto, no sería admisible el relegar la deseabilidad social a un mero factor de error sino que ha de ser considerado como un indicador de la adaptación social del sujeto (Bardburn y Sudman, 1979) Sin embargo, la varianza explicada por la deseabilidad social es sustancialmente inferior a la explicada por el contenido del test. Entre otros, Fierro, 1982 en Fernández-Ballesteros, 1992)

SIMULACION

La simulación es el falseamiento de las respuestas deliberado de parte del sujeto y está muy en función del objetivo del examen; es el engaño intencional para aparecer "mejor" o "peor". De acuerdo con Anastasi (1971), la distorsión de las respuestas es en gran medida de tipo motivacional. El engaño o simulación se realiza con un propósito, con un motivo que se alcanzaría hipotéticamente con ciertos resultados del test. Como se expuso al inicio de este capítulo, la situación influye en la distorsión y ésta se presenta frecuentemente en evaluación de personal. Lanyon y Goodstein (1977) agregan que pueden combinarse diferentes tipos de distorsión. De tal forma, la "defensividad podría referirse tanto al conjunto de respuestas de revelar en general muy poco acerca de uno mismo como al estilo de respuesta de elegir la categoría de "no puedo decir" o "no estoy seguro" en los cuestionarios que incluyeran esta opción.

¿Existe una relación entre la capacidad de simulación y otras variables de personalidad? Empleando grupos de solicitantes de empleo y alcohólicos, Kanter (1963, citado por Lanyon y Goodstein, 1977) encontró que los sujetos mejor ajustados eran los más exitosos en producir perfiles distorsionados para "aparecer mejor" en el Inventario Psicológico de California.

Un estudio de Lanyon (1967b en Lanyon y Goodstein, 1977) ha sugerido que estos descubrimientos pudieran asimismo ser válidos en la simulación de perfiles patológicos.

En este estudio, estudiantes universitarios bien ajustados se mostraron superiores a estudiantes mal ajustados en su capacidad para simular un patrón MMPI sugestivo de personalidad psicopática. Megargee (1971) ha sugerido que el falseamiento y la capacidad para hacerlo pueden ser una variable de personalidad de considerable importancia y no solamente un factor indeseable que tenga que corregirse.

Por otro lado, hay considerable evidencia para sugerir que los individuos perturbados, especialmente a los pacientes psiquiátricos, no les es posible simular completamente un perfil normal. Grayson y Olinger (1957, en Lanyon y Goodstein, 1977) demostraron que sólo el 11% de un grupo de pacientes psiquiátricos pudo aparentar un perfil normal del MMPI.

En situaciones laborales, es marcada la tendencia a la simulación favorable, para aparecer "mejor". Guierasch (1972 en Thornton y Guierasch 1980) refiere que en un estudio empírico de 34 instrumentos usados en evaluación industrial, 20 de ellos mostraron que el falseamiento incrementa las respuestas favorables, uno no mostró efectos de falseamiento y el resto mostró resultados ambiguos. Thornton y Gierasch (1980) aplicaron una prueba de selección construida empíricamente a dos grupos, uno de ellos bajo instrucciones de simular el ser solicitantes de empleo muy motivados y el otro bajo las instrucciones usuales. Encontraron que el primer grupo obtuvo perfiles "inflados" con incrementos significativos en siete de las diez escalas en el sentido que incrementaría la probabilidad de ser contratado.

MÉTODOS PARA REDUCIR LA DISTORSIÓN

Existen formas de reducir los efectos de los diferentes tipos de distorsión de las respuestas, algunos de los cuales se van a exponer a continuación.

Para reducir los ESTILOS DE RESPUESTA como el asentimiento.

Una forma mencionada por Fernández-Ballesteros (1992) entre otros, para reducir el asentimiento, consiste en invertir el significado del contenido del reactivo o la pregunta se formula negativamente, es decir, se equilibra la direccionalidad de los elementos en la totalidad del inventario. Esto tiene la limitación de que se dificulta su comprensión. Rorer y Goldber, (1965, en Fernández-Ballesteros, 1992) en una amplia recopilación concluyeron que el asentimiento no explica una parte significativamente relevante de la varianza en las respuestas a los autoinformes de alternativa dicotómica.

En el formato escalar, existen procedimientos estadísticos para el análisis y el control de las respuestas intermedias o extremas inusualmente numerosas de parte de un sujeto. (Fernández-Ballesteros, 1992) P 248 a 250

Block (1965 en Lanyon y Goodstein, 1977) elaboró escalas libres de condescendencia para el MMPI modificando los reactivos hasta que estuvieran equilibrados para su identificación como "cierto" y "falso". Una serie de análisis factoriales de las respuestas a estas escalas modificadas (y presuntamente libres de condescendencia) produjo la misma estructura factorial que había sido obtenida con las escalas originales, llevando a Block a concluir que la condescendencia no era una variable significativa en la estructura factorial del MMPI y que esta era determinada, en consecuencia, por el contenido del reactivo.

Para reducir la DESEABILIDAD SOCIAL, Edwards (1957 en Catell y Kline, 1982) encontró que en elementos del MMPI, había una correlación sustancial entre la deseabilidad social que los jueces asignaran a un elemento y el porcentaje de respuestas que ese elemento recibía. Edwards concluyó que gran parte de la varianza de los elementos de los cuestionarios reflejaba sencillamente las diferencias en deseabilidad social y elaboró una prueba con elementos equilibrados en cuanto a su deseabilidad social usando un método de jueces.

La conclusión de que el conjunto de respuestas de deseabilidad social es una variable básica que influye en las respuestas de los inventarios de manera importante, ha sido cuestionada seriamente Norman, (1967 en Lanyon y Goodstein, 1977) por ejemplo, ha observado que los diseños estadísticos empleados por Edwards, y por ende sus resultados, "son básicamente irrelevantes" para el problema en consideración. Otra demostración empírica de que la deseabilidad social no es un factor distorsionante en las respuestas a los reactivos de inventarios de la personalidad, fue referida por Hellbrun y Goodstein, (1961a, en Lanyon y Goodstein, 1977) que realizaron un estudio sobre la base de que la deseabilidad social es diferente de individuo a individuo y con esto cuestionaron el método de jueces seguido por Edwards para homologar el grado de deseabilidad social de sus reactivos.

Catell y Kline (1982) mencionan que la mayoría de los constructores de pruebas consideran que la influencia de la deseabilidad social en los elementos de la prueba puede ser disminuida al máximo al redactar los elementos. El impacto de la deseabilidad social es diferencial dependiendo del factor que se mida.

Para contrarrestar la SIMULACION deliberada, dos son los procedimientos generalmente utilizados: 1. Aprovechando del proceso de aplicación y solicitando específicamente la sinceridad y honestidad al contestar la prueba.

2. La adición al inventario de escalas especialmente diseñadas para identificar la tendencia a falsear las respuestas en uno u otro sentido (Fernandez-Ballesteros, 1992), llamadas escalas de validez, de mentira o sinceridad, etc. Es conveniente incluir aquí lo referido por Norman (1963 en Lanyon y Goodstein, 1977) en relación a las escalas de validez del MMPI pero que puede generalizarse a cualquier inventario: estas escalas de validez son útiles en la medida que la validez predictiva de las escalas originales sea bastante alta en primer lugar.

3. 3 LAS ESCALAS DE VALIDEZ.

Los inventarios de personalidad cuentan con escalas adicionales que identifican diferentes formas de simulación y algunos de estilos de respuestas. Se construyen ex-profeso para cada inventario; posteriormente, en este capítulo se referirán las del MMPI y las del 16 PF. Otros inventarios que cuentan con sus escalas de validez son el de Comrey, 1970 y el de Jackson, (1984) referidos por Christiansen et al (1994)

El Inventario Psicológico de California de Gough, cuenta con tres escalas de validez, la de Bienestar, Buena impresión y Comunidad. Las primeras dos se elaboraron a partir de los reactivos que propiciaban determinadas respuestas de parte de personas normales a quienes se les pedía que falsearan para aparecer peor (escala de bienestar) o para aparecer mejor (escala de buena impresión). La escala de comunidad es un recuento de las respuestas más populares. Esta prueba se derivó del MMPI pero, a diferencia de éste que se dirige a evaluar los trastornos psiquiátricos y de desajuste, el CPI se enfoca a aspectos de la personalidad normal. (Aiken, 1995)

Otra técnica de control del falseamiento deliberado, mencionada por Lanyon y Goodstein, (1977) es el uso de reactivos sutiles; consiste en elegir reactivos con validez predictiva o concurrente empíricamente demostrable pero de escasa o nula validez aparente. Esto tiene la limitación de que los reactivos sutiles tienden a ser menos válidos o útiles que los reactivos más obvios. Agregan estos autores que Norman (1963), quien desarrolló un complejo método para controlar la simulación al eliminar cuidadosamente los reactivos más obvios en sus escalas, llegó a la conclusión de que el uso de tal procedimiento no puede generalizarse ampliamente.

Cabe aquí mencionar el planteamiento de Wallace (1966,1967 en Lanyon y Goodstein, 1977). Este investigador propone considerar a la evaluación de la personalidad o sus dimensiones como una prueba de ejecución máxima en lugar de una prueba de ejecución habitual, es decir, de conductas típicas. De tal modo, una medida de dominancia, por ejemplo, pudiera involucrar una prueba situacional en la que se requeriría del sujeto que respondiera tan dominantemente como le fuera posible y sería evaluado en su actuación real al confrontar estas expectativas de tarea. Para una prueba de ejecución máxima, mas que de ejecución habitual, el problema del engaño o de la defensividad viene a ser un tanto irrelevante. En relación a esto, debe enfatizarse que la distorsión o simulación existe principalmente en los inventarios de personalidad por demandar respuestas verbales y de autodescripción de parte del evaluado; los inventarios no evalúan conductas directamente.

Esta alternativa propuesta por Wallace se ubica más bien dentro de las técnicas multidimensionales como el simulacro Assessment Center en el cual se evalúan algunos rasgos de la personalidad a través de la observación -por parte de otros- de conductas en situaciones simuladas y efectivamente, la probabilidad de distorsión es muy baja.

3.3.1 LOS TESTS DE INTEGRIDAD

Existe toda una constelación de instrumentos desarrolladas en EE.UU. cuyo origen fue la preocupación de las organizaciones -sobre todo laborales- de protegerse contra personas de alto riesgo que puedan cometer robos, acciones deshonestas o cualquier acto que atente contra la seguridad y los intereses de la organización. Estos instrumentos se llaman genéricamente tests de integridad (*integrity tests*) y se utilizan sobre todo con propósitos de selección de personal. El polígrafo o detector de mentiras combinado con cuestionarios o entrevistas, fue utilizado en EEUU para identificar posibles ladrones y mentirosos hasta que fue prohibido en 1988 (Ailken, 1996). Paul Sackett y James Wanek (1996) han realizado una amplia y documentada revisión del tema la cual aborda aspectos legales, conceptuales, compilación de investigación y la relación de estos instrumentos con el enfoque de los "cinco grandes" factores de personalidad.

Sackett et al (1989 en Sackett y Wanek, 1996) refieren que estos instrumentos pueden clasificarse genéricamente en dos categorías

1. los tests "abiertos" de integridad (*overt integrity tests*) cuyo propósito es evidente para el respondiente, ej el *London House PSI*, *The Reid Report* y el *Stanton Survey*.
2. Las medidas orientadas a la personalidad (*personality oriented measures*) que son más similares a las pruebas de personalidad o son pruebas de personalidad como el Inventario Psicológico de California que no evalúa específicamente la conducta deshonestas pero contienen algunas escalas para apreciarla. Ejemplos: *Personnel Reaction Blank*, El *PDI Employment Inventory* y la Escala de Veracidad de la serie de personalidad de Hogan.

Estos mismos autores mencionan que se han encontrado correlaciones substanciales entre los resultados de los tests de integridad y las escalas de mentira o de deseabilidad social. Un ejemplo es el estudio de Cunningham (1989 en Sackett y Wanek, 1996) quien asocia los puntajes de escalas de falseamiento con los de tests de integridad. El construyó escalas de "integridad motivacional" para evaluar estrategias de autodescripción tales como el engaño, las respuestas al azar y la legitimación de la deshonestidad (la deshonestidad como norma) y las aplicó, junto con el Reid Report, a una muestra de solicitantes de empleo. Encontró que los sujetos clasificados por el Reid Report como "no recomendables para la contratación" tenían más alta probabilidad de obtener puntajes altos en las tres escalas mencionadas.

En los tests de integridad también se han cambiado experimentalmente las instrucciones para inducir el falseamiento y en general se ha encontrado que, bajo estas condiciones, las personas pueden aumentar sus puntajes en los tests de integridad para que los favorezcan. (Dean, 1990, Lobello y Sims, 1993, Moore, 1990 y el referido Cunningham, et al, 1994, en Sackett y Wannek, 1996)

Se ha discutido e investigado bastante si los tests de integridad abiertos son más fáciles de distorsionar, por lo manifiesto de su propósito, que las medidas orientadas a la personalidad, pero los resultados han sido ambiguos. En resumen, crece la evidencia de que, estudiantes en condiciones experimentales, pueden aumentar sus puntajes en tests de integridad cuando reciben instrucciones de simular el ser un candidato a empleo y, por ende, presentar su mejor imagen, pero esto no afecta la validez de criterio de esos tests, estudiados con muestras de solicitantes de empleo en escenarios aplicados, como lo refieren Ones et al (1993b en Sackett y Wannek, 1996) con base en resultados de sus meta-análisis.

La latencia de la respuesta

Un enfoque muy diferente y que abre posibilidades promisorias en la investigación del falseamiento de respuestas es el estudio de tiempos de respuesta. Este fue empleado por McDaniel y Timm, (1991 en Sackett y Wannek, 1996), para detectar el falseamiento en los tests de integridad. Ellos aplicaron una batería de tests a una muestra de estudiantes universitarios y utilizaron una computadora para medir el tiempo de respuesta a los ítems. La batería incluyó la *Hogan Reliability Scale*, una escala de deseabilidad social y un subconjunto de reactivos del *Inweid Selection Inventory* relacionados con actividad criminal y uso de drogas.

Los estudiantes primero contestaron la batería simulando ser un solicitante de empleo y en seguida la contestaron honestamente. Se consideró que aquellos que cambiaron sus respuestas, habían contestado deshonestamente la primera vez. Estos se tomaron 0,66 seg. mas de tiempo de respuesta, en promedio. Las diferencias en la latencia de la respuesta fueron moderadas por el contenido del ítem y las mayores diferencias se encontraron en los contenidos de uso ilegal de drogas y preguntas relacionadas con empleo. Sobre esta misma línea de investigación está el trabajo de Siem (1991 en Colom, 1995) que estudio la validez predictiva de las latencias de respuesta en pruebas de personalidad administradas por computadora aunque no precisamente en lo relacionado a falseamiento. Colom enfatiza la importancia de las latencias de respuestas como variable a considerar en la medida de las dimensiones de la personalidad en general.

3.3.2 LAS ESCALAS DE VALIDEZ DEL MMPI.

El MMPI original cuenta con cuatro indicadores de validez: el número de respuestas omitidas y tres escalas de validez especialmente diseñadas para tal fin: la escala L, la escala F y la escala K.

La interpretación de las escalas de validez indica básicamente la actitud del sujeto al responder al inventario. La relación entre los puntajes de las tres escalas evidencia la actitud de la persona hacia su propia descripción. En su libro de Interpretación del MMPI, Rivera (1991) incluye 13 combinaciones posibles de las tres escalas de validez sobre la base de los puntajes T, bajos, medios y altos. Esta autora ofrece una interpretación de cada una de las combinaciones. Por su parte, Nuñez (1994) presenta las nueve combinaciones de las escalas de validez investigadas por Lachar (1974) con sus correspondientes interpretaciones.

Para que una prueba pueda considerarse como no-válida, Rivera (1991) menciona tres condiciones indispensables: a) Una puntuación en bruto de la escala F de 20 o más. b) El índice de Gough (F menos K , puntajes brutos) debe ser de 9 o más. c) ninguna de las diez escalas clínicas se debe encontrar en la zona de "normalidad estadística" T 55 a T 45. Rivera no menciona un número máximo de respuestas ausentes; si éste es muy alto, no puede considerarse un inventario válido.

ESCALA DE FRASES OMITIDAS

Las ausencias de respuesta se llama también Escala de frases omitidas y se identifica con una interrogación: ?. Varias son las situaciones que provocan un alto número de respuestas omitidas: una pobre comprensión de los reactivos, ciertos tipos de perturbación emocional como depresión o paranoia, tendencias muy marcadas a la intelectualización, etc. (Nuñez 1994) El número de respuestas omitidas indica, en sujetos normales, una posición defensiva al responder. (Aiken, 1990). Mehl (1960 en Nuñez, 1994) propone que el máximo permisible de reactivos sin contestar debe ser de 30.

En mi práctica profesional en el campo de la evaluación de personal, he encontrado un número alto de omisiones al MMPI en casos de personas que se vieron muy forzadas a contestar el cuestionario, en personas con pobre comprensión de la lectura y en casos de personal de niveles directivos que percibieron la prueba como "muy clínica" e inapropiada para evaluar personas con propósitos laborales.

LA ESCALA L (DE MENTIRA)

Indica el grado de franqueza del sujeto al contestar, la tendencia a cubrir faltas socialmente inaceptables pero también puede evidenciar un intento deliberado por distorsionar las respuestas. Se asocia con una capacidad intelectual baja y con limitaciones culturales (Núñez, 1994) Rivera, (1991) por su parte, considera la escala L como indicador de necesidad de aceptación social.

La Escala L es una escala diseñada para evaluar la alteración de las respuestas cuando éstas reflejan las actitudes convencionales con que el individuo pretende mostrarse apegado a los valores sociales, como alguien que reacciona de acuerdo con las expectativas del grupo. Mientras mayor es la elevación en esta escala, mayor es la necesidad del individuo de mostrarse socialmente ajustado aun a costa de su propia individualidad. Los puntajes bajos indican intentos de demostrar el ser rígido y estricto aunque no este muy integrado al grupo (Rivera, 1991)

LA ESCALA F.

La escala F, permite localizar los efectos de que el sujeto consteste a los reactivos de manera descuidada o confusa, ya que se autodescribira como alguien que tiene características raras o improbables. (Mischel 1988)

La escala F del MMPI, que representa un esfuerzo por identificar a sujetos que están contestando de una manera atípica, contiene reactivos que rara vez son respondidos en cierta dirección particular, debido a que las respuestas raras, por lo común son también socialmente no deseables o psicopatológicas. La escala F, sirve asimismo para identificar a las personas que están tratando de aparecer como "malas". Existe una buena cantidad de evidencias de investigaciones que muestran que las escalas de validez del MMPI son hasta cierto grado eficaces para identificar la simulación, por ejemplo, Gough (1947) Kofer, Chabce y Judson, (1949) citados por Lanyon y Goodstein, (1977)

Esta escala F se formo con reactivos de contenido que plantea alteraciones psicológicas severas. Una elevación indica una fuerte tensión interna y la tendencia a la exageración de la propia problemática, frecuentemente ocasionada por la sensación de impotencia para manejarla por si mismo. Esta escala puede indicar también una falta de comprensión del contenido de los reactivos, en especial, de los que están expresados en forma negativa. En los casos de puntajes significativamente bajos, puede asociarse con una reducción de la capacidad de insight (Rivera, 1991)

Dalshtrom (en Núñez, 1994) menciona que una elevación en esta escala también puede indicar falta de comprensión de la lectura de los reactivos o una contestación al azar.

LA ESCALA K

La Escala K está compuesta por reactivos que se refieren a la tendencia a evitar el reconocer la propia problemática y es fundamentalmente en este sentido que debe ser interpretada. Esta escala contribuye como factor de corrección de la prueba al elevar los puntajes de cinco escalas clínicas: 1, 4, 7, 8 y 9, que son las que tienen mayor probabilidad de alterarse por ser las más sensibles a dicha tendencia. (Rivera, 1991)

Una K elevada, agrega Rivera (1991), se interpreta con la exageración de una actitud de reserva en la autodescripción con la que la persona intenta mostrarse capaz de enfrentar y resolver sus problemas por sí misma. Con frecuencia, este tipo de personas rechazan las sugerencias terapéuticas e incluso dificultan la entrevista al restar importancia a sus conflictos. Una K reducida indica que la persona se siente incapaz de enfrentar y resolver sus problemas por sí misma. K es asimismo sensible a la situación. En los casos en que se responde al MMPI bajo cierta presión, se observa una tendencia importante a incrementar este puntaje. Cuando las personas desean contestar la prueba por propia iniciativa, por ej. al buscar ayuda psicológica, es más probable que los puntajes en esta escala se reduzcan.

La escala K del MMPI elaborada empíricamente para corregir los aspectos más sutiles de la defensividad de la situación de examen, es un buen ejemplo de un instrumento de corrección. La escala K fue elaborada al comparar las respuestas de personas normales con las de pacientes psiquiátricos cuyas puntuaciones en las escalas clínicas se encontraban en la gama normal, y de quienes se podría suponer que estaban tratando de aparecer como "buenas". La puntuación bruta de la escala K se usa directamente como una variable correctiva y se suma en fracciones diferentes a las puntuaciones en cinco de las escalas clínicas. Se ha demostrado que el uso de la escala K en esta forma aumenta el poder discriminatorio de dichas escalas, particularmente en la gama crucial de la parte media de los valores de puntuación. (Dahlstrom y Welsh, 1960, en Lanyon y Goodstein, 1977)

Debe notarse que las puntuaciones en las escalas como la K tienen valor interpretativo en sí mismas, en especial en poblaciones normales, por ejemplo, tanto Smith (1959) como Heilbrun (1961) citados por Lanyon y Goodstein, (1977), han logrado demostrar que, en una población normal de estudiantes universitarios, la escala K es tenida más por un índice de salud mental o de ajuste que por una medida de defensividad.

El índice F-K de Gough. Este identifica a aquellas personas que distorsionan el inventario simulando un cuadro patológico. Se obtiene al restar la puntuación natural de K de la puntuación natural de la escala F. Si la diferencia entre las dos puntuaciones es de 9 o mayor, puede sospecharse que se esté fingiendo un trastorno emocional. Lachar, (en Nuñez, 1994), presenta las interpretaciones de diferentes rangos de este índice.

La versión revisada del MMPI, el MMPI-2, conserva las escalas de validez originales y agrega tres adicionales: la Fp, infrecuencia posterior, la INVAR, inconsistencia de las respuestas variables y la INVER, inconsistencia de las respuestas verdaderas. Estas complementan la información sobre la forma en que fue contestada la prueba. (Catálogo 1997 de pruebas de El Manual Moderno)

La escala Fp (infrecuencia de la página posterior, indica la tendencia del sujeto a contestar al azar en los últimos reactivos. Proporciona una verificación adicional útil sobre todo para las escalas suplementarias y de contenido. Las escalas INVER (inconsistencia de las respuestas verdaderas) e INVAR (inconsistencia de las respuestas variables) son escalas de validez nuevas del MMPI-2; proporcionan un índice de la tendencia a responder de manera inconsistente o contradictoria. La Inver identifica aquellos reactivos que se oponen en contenido y a los cuales el sujeto tendría que responder de manera distinta. La Invar detecta contestaciones al azar, complementando la interpretación de la escala F. Si la F y la Invar son altas, indican respuestas al azar pero si la F es alta y la Invar es baja, puede tratarse de sujetos mentalmente perturbados o que intentan aparecer más perturbados de lo que realmente están. (Lucio 1995 en Valencia, 1996)

El estudio de las escalas de validez del MMPI y los diferentes tipos de falseamiento de respuestas ha dado lugar a múltiples investigaciones que se iniciaron en escenarios clínicos para identificar a las personas que estaban tratando de fingir o exagerar síntomas. Este ha sido el método psicométrico más ampliamente usado para estudiar el engaño o el fraude psicológico (Adelman y Howard, 1984, Rogers, 1984, Ziskin, 1984, en Valencia, 1996) que posteriormente cobró enorme importancia en EEUU en los casos legales en que el fingimiento podía evitar sentencias o permitía ganar beneficios económicos como pensiones o indemnizaciones. (Valencia, 1996)

Grow, Mcvaugh y Eno (1980 en Valencia, 1996) evaluaron la efectividad de 13 diferentes técnicas para detectar a los sujetos simuladores con el MMPI original. Consideraron las 3 escalas de validez, la escala de reactivos sutiles-obvios de Wiener (1948) y el índice de Gough entre otros. Aplicaron MMPI a dos muestras de sujetos, estudiantes universitarios y pacientes psiquiátricos, en tres situaciones, simular estar mal adaptados, simular estar bien adaptados y de sinceridad. Entre los principales resultados está la confirmación del valor de la escala F y el índice de Gough para identificar la simulación en ambas direcciones de las dos muestras. Una $F > 15$ e $I.Gough > 7$ identificaron al 80% y 79% de los estudiantes que fingieron estar mal mientras que una $F > 15$ e $I.Gough > 7$, tuvieron cada uno una efectividad de 70% para identificar a los pacientes que simulaban estar mal. El 60% de los pacientes simulando estar bien fueron detectados con $I.Gough < -11$ pero este mismo índice sólo identificó al 30% de los estudiantes fingiendo estar bien.

Graham, Wats y Timbrook, (1991 en Valencia, 1996) informan de un estudio similar con el MMPI-2 que fue aplicado también en condiciones de falseamiento artificial positivo o negativo en dos muestras, una de estudiantes universitarios y otra de pacientes psiquiátricos. En la condición de falseamiento negativo de los estudiantes, la escala F y todas las escalas clínicas tienden a elevarse significativamente más que en la muestra de pacientes. Se encontró que la escala K es inadecuada para diferenciar perfiles de sujetos que simulaban estar bien de los honestos pero la escala L sí es efectiva para detectar a los sujetos que simulan estar bien adaptados.

Un estudio realizado con el MMPI-2 en México (Valencia, 1996) que analizó la efectividad de las escalas de validez para diferenciar sujetos simuladores de sujetos honestos, obtuvo resultados similares a los de Grow et al. y de Graham, et al. recién expuestos. Se consideraron dos muestras, de estudiantes universitarios y de pacientes psiquiátricos, clasificadas por sexo. El MMPI-2 se aplicó en condiciones normales y de falseamiento inducido, tanto positivo y negativo. Las escalas que resultaron más exitosas en la detección de los sujetos que simularon estar mal, fueron las escalas F, Fp y el índice F-K mientras que la escala L resultó ser la más exitosa para identificar a los sujetos que simulaban estar bien. En este estudio, como en los dos que sirven de referencia, fue más fácil identificar a los sujetos que fingían estar mal que a los que fingían estar bien.

3.3.3 LAS ESCALAS DE VALIDEZ DEL 16 PF-A

El cuestionario 16 Factores de personalidad es sensible al falseamiento común a todos los inventarios. Aunque sus autores sostienen que sus elementos están balanceados y que ambos polos de los factores son igualmente atractivos para reducir el falseamiento, Irvine y Gendreau (1978) encontraron claras evidencias de sensibilidad al falseamiento en la mayoría de las escalas. Estos investigadores aplicaron el 16 PF a dos grupos, uno de prisioneros y otro de estudiantes universitarios, consideraron las 16 escalas primarias y dos de las secundarias. Subdividieron los grupos en tres, uno contestó el test con instrucciones normales, otro con instrucciones de "aparecer mejor" y el tercero lo respondió con instrucciones de "aparecer peor". Encontraron que solamente seis de los 18 factores evaluados, resistieron el falseamiento para "aparecer mejor" y solo dos de los 18 factores soportaron el falseamiento para aparecer "peor".

Seisdedos (1985) menciona los principales tipos de distorsión del Cuestionario 16 PF y la forma en que se ha tratado de reducirlos

a) Estilos de respuesta. En este punto se va a revisar la condescendencia, las respuestas intermedias y el sabotaje en el 16 PF.

La condescendencia. Las distintas formas del 16 PF fueron construidas procurando llegar a un equilibrio de los elementos más susceptibles a la condescendencia. Para ello se seleccionaron los elementos con saturaciones significativas en el factor deseado y con correlaciones mínimas respecto al criterio de condescendencia, (estimado a partir del recuento de frecuencias asentidas, así como de un análisis formal de los contenidos) Sin embargo, la eliminación de la condescendencia no ha sido total. (Seisdedos, 1985) El 16 PF-5 incluye una escala de Aquiescencia o condescendencia.

En estudios sobre todo con el MMPI, se encontró que una limitación en la comprensión verbal tiende a aumentar la condescendencia (Seisdedos, 1985); en este sentido, en mi experiencia con el 16 PF he encontrado que la Escala B puede ser útil como un indicador grueso del nivel de comprensión de la lectura; si la calificación es baja debe investigarse si esta limitación afecta la comprensión de los reactivos y la manera en que se manifiesta; con un incremento de la condescendencia, de las respuestas al azar, etc.

McGe encontró una tendencia a la condescendencia en diversas escalas cuando el tiempo de aplicación se reducía al mínimo. (1962 en Seisdedos, 1985) por lo que se deduce que una reducción del tiempo previsto de aplicación puede probablemente aumentar la condescendencia.

Otra estilo de respuestas que he identificado en mi experiencia con este instrumento en evaluación de personal, es la presencia de un número inusual de respuestas intermedias en un intento del evaluado por evitar dar información de sí mismo. No se ha establecido punto de corte para considerar esta distorsión como crítica o invalidar la prueba; esto debe hacerse sobre bases empíricas. Este sesgo puede reducirse al impartir las instrucciones relejendo el párrafo que se refiere a evitar estas respuestas "de indecisión o de duda"

Una variación de este tipo de falseamiento son los protocolos con ausencia de respuestas o con respuestas dobles. Estos deben identificarse antes de finalizar el proceso de aplicación para solicitar al sujeto que las conteste. Asimismo, las instrucciones deben colaborar para evitar este estilo de respuestas.

El sabotaje deliberado del sujeto que rehusa cooperar a las instrucciones dadas es otro estilo de respuesta. Se identifica por una relación inusual en las respuestas, una irregularidad en los intervalos o una regularidad espacial que no tiene significado psicológico, por ejemplo, que todas las respuestas caigan del lado derecho de la hoja de respuestas, una alternancia izquierda-derecha, etc. El Manual del 16 PF (Catell, Eber y Tatsuoka, 1970-1985) proporciona una fórmula para calcular este tipo de falseamiento que denomina sabotaje la cual da como resultado un índice de desacuerdo entre las dos mitades de los puntajes de cada escala. Este es un procedimiento más bien complejo y debe utilizarse solamente cuando la situación justifique. (p 50)

Seisdedos (1985) amplía la explicación del cálculo del sabotaje (ver p 28) pero recomienda que solo se use cuando se han aplicado dos formas del 16 PF. Menciona asimismo que se considera sabotaje la presencia de un número excesivo de estenes extremos, 1 o 10 y establece que más de tres de estos puntajes se consideraría atípico, según estudios españoles. Por los estudios de O'Dell (1971 en Seisdedos, 1985) se deduce que la distorsión por azar es distinta a la distorsión por sabotaje.

b) La deseabilidad social en el 16 PF. Desde su construcción se buscó un equilibrio en los elementos en cuanto a deseabilidad social. Además, en este cuestionario existe una escala, la Q3, que apunta precisamente al nivel de control de la propia imagen social, que podría hacer innecesaria una escala complementaria de deseabilidad social. Por otro lado, no está del todo claro que esta tendencia de los sujetos pueda considerarse como un factor único pues diversos análisis factoriales han mostrado la existencia de un patrón complejo y muy relacionado con el rol que el sujeto está desempeñando en la situación de examen. (Seisdedos, 1985) El 16 PF-5 incluye una escala de deseabilidad social denominada IM.

c) La simulación. Es factible aminorar el efecto de la simulación por medio de las instrucciones pero el psicólogo no debe perder de vista el efecto de la situación. Las escalas de validez identifican el tipo y el grado de distorsión.

Seisdedos (1985) refiere un estudio realizado en España con 345 varones y 530 mujeres que habían contestado al MMPI y al 16 PF-A en una situación de selección: el ingreso a la universidad. Se dicotomizaron las muestras por la mediana en la escala 'F-K' de disimulación o índice de Gough (1950) y tanto los varones como las mujeres que distorsionaban aumentaban sus puntuaciones en las siguientes escalas del 16 PF: A, C, G, H, M, N y Q3. (cuyo conjunto apunta hacia extroversión) y disminuyeron las de L, O y Q4. (que inhiben la manifestación de ansiedad). En estos análisis, el 74% de las correlaciones interescalas entre los varones y el 50% entre las mujeres estaban influidas por la distorsión.

Karson y O'Dell (1989), consideran que uno de los principales problemas en el examen de la personalidad es el hecho de que las personas que contestan a los cuestionarios intentan presentar una imagen falsa de sí mismos y que ésta puede deberse a una simulación directa, a una baja atención al contestar o a otras causas. En su procedimiento para interpretar el perfil, presentan como primer paso el considerar el puntaje de las escalas de validez, distorsión motivacional DM, mala imagen y azar y en el informe lo primero que se menciona es la puntuación de la Distorsión motivacional DM.

Karson y O'Dell (1989) identifican tres tipos de falseamiento en el 16 PF: las personas pueden intentar presentar una imagen mejorada de sí mismos lo cual ha sido llamado en la literatura del 16 PF Distorsión Motivacional, abreviándose por sus siglas, DM. También existen los intentos de aparecer "peor" que la realidad, es la Distorsión Negativa. Otro caso se presenta cuando la persona no le da la suficiente importancia al cuestionario y contesta descuidadamente o al azar, la escala para identificar esta tendencia se llama Escala de Azar.

Karson y O'Dell (1989) recomiendan que el puntaje de las escalas de validez es particularmente importante en las situaciones en las que el examinador no tiene ocasión de ver personalmente al sujeto y solo interpreta los resultados de tests aplicados por otra persona.

LA ESCALA DE DISTORSION MOTIVACIONAL DM EN EL 16 PF-A.

Anastasi (1971) al referirse al falseamiento de respuestas en tests de personalidad menciona que la distorsión de las respuestas es en gran medida de tipo motivacional. Considera que las instrucciones por sí solas deberían eliminar los intentos por falsear las respuestas pero esto no es factible en ciertas situaciones. Las personas inadaptadas, por ejemplo, responderán al test de una manera demasiado emocional para seguir un argumento racional. Existen otras situaciones en que los objetivos del examinador y del examinado son fundamentalmente opuestos y esto provoca el falseamiento. Al combinarse una fuerte motivación por conseguir cierto objetivo con un sentimiento de inseguridad respecto al resultado de un test, es muy probable que se intente falsear las respuestas. Esto se presenta de manera muy clara en los casos en que la persona evaluada está concursando por un empleo: sabe que los resultados serán tomados como base para la contratación y por supuesto, sus respuestas serán encaminadas a presentar la mejor imagen de sí mismo.

Se ha llamado Distorsión Motivacional DM al tipo de falseamiento de respuestas en que el evaluado intenta aparecer mejor. Este tipo de falseamiento es el más frecuente al evaluar personal. Catell construyó una escala de distorsión para las formas C y D que originalmente estaban diseñadas para evaluar personal por ser bastante más cortas pero las forma A y B no cuentan con escalas de validez desarrolladas directamente por Catell (**Karson y O'Dell, 1989**).

Específicamente las personas que solicitan trabajo o ingreso a escuelas están motivadas a distorsionar sus resultados para ser vistos de la mejor manera posible. La cuestión de la distorsión motivacional en las pruebas psicológicas es de máxima importancia para el psicólogo. (Winder, O'Dell y Karson, 1975)

Es posible estimar la distorsión motivacional DM si las pruebas se hacen en más de una sesión y en situaciones distintas, por ejemplo, en selección de personal y en consejo psicológico, desmenuzando por partes la personalidad, el papel y los factores cognoscitivos estereotipados, como demostró Krug (1970 en Catell y Kline, 1982) quien aplicó la prueba 16 PF en cuatro condiciones diferentes: anónimo, búsqueda de trabajo, presentación de sí mismo en términos estéticos e ideales y en una mera obtención de datos. Las pruebas T evaluaron además la fuerza del interés por obtener un trabajo y el interés hacia los miembros del sexo opuesto -factores del papel social. Los resultados mostraron un factor del papel social en ambas situaciones, que afectaba a las puntuaciones en los demás factores. De todas formas, los pesos de los factores del papel social no fueron superiores al promedio de los factores generales de la personalidad. A lo largo de todos los rasgos y de las cuatro situaciones indicadas, la adscripción de los efectos perceptuales de los restantes rasgos, aumentó significativamente la validez de las puntuaciones factoriales, que se incrementaron luego por la adición de los pesos del papel social desmenuzando así las distorsiones motivadoras. (Catell y Kline, 1982)

Otra forma de aproximarse a la distorsión motivacional DM está en función de la teoría de la perspectiva del rasgo, mencionada anteriormente y que enfatiza la influencia del observador al evaluar la conducta, sea la propia o sea la de otros. Catell, Eber y Tatsuoka (1970-1988), mencionan que este enfoque evita el análisis de la distorsión por partes y en su lugar postula un modelo general. Explican que las respuestas distorsionadas, como cualquier otra conducta, están en función de la personalidad del observador y de la situación. Estos autores proponen una ecuación para evaluar la distorsión que considera el rasgo, el rol del sujeto (si es solicitante de empleo, paciente, etc.), la situación y una constante C que es el promedio de la Distorsión en esa situación. (p 51) Esta forma de corregir la distorsión sirve para la Distorsión motivacional DM o la Deseabilidad Social, DS pero presenta ciertos problemas prácticos pues se necesitan los valores C para cada situación particular y, en el caso de la deseabilidad social, habría esperar una variación mayor, por ejemplo, entre sexos, entre culturas, etc. por lo que se necesitarían datos de C sistemáticamente actualizados

Seisdedos, (1985) refiere un estudio similar al de Winder para derivar una escala de distorsión motivacional DM y de Azar para el 16 PF-A, realizado en España. Se conformaron las escalas, se revalidaron en una muestra mayor, se obtuvieron baremos y se realizó un análisis correlacional con las 16 escalas primarias. Este procedimiento de la construcción de la escala de distorsión motivacional DM se detalla en el capítulo siguiente en tanto que el estudio de la escala de azar se muestra en el siguiente punto.

ESCALA DE AZAR DEL 16 PF-A. Pretende discriminar a los sujetos negativos o poco cooperativos que desatendiendo las instrucciones contestan al azar o de modo infrecuente. Se le llama también escala de negación pero no debe confundirse con la escala de distorsión negativa o DN. Esta escala se construyó a partir de la localización de las respuestas infrecuentes en las condiciones usuales de aplicación, similarmente a la Escala F del MMPI. Esta escala podría llamarse, como un nombre alternativo al de negación, escala de "renuencia"

Es interesante observar los efectos de los distintos tipos de distorsión en los perfiles resultantes. Unas contestaciones dadas al azar conducen a un aplanamiento en la forma del perfil a causa del modo equilibrado en que se construyeron los elementos y las alternativas de respuestas. Esto ocurre así en todas las escalas excepto en la B, inteligencia, en la que la puntuación descenderá bastante. (Karson y O'Dell, 1989)

En España se ha construido una escala para identificar las respuestas al azar al 16 PF-A (Seisdedos, 1985) paralelamente a la derivación de la escala de distorsión motivacional DM y con las mismas muestras. Se tomaron únicamente a los 247 sujetos que en una primera fase de construcción de la Escala DM, habían contestado sinceramente al 16 PF-A. Se seleccionaron los 22 elementos que tenían una alternativa infrecuente. Entre ellos, en cuanto a su pertenencia a las escalas, hay tres en A, seis en B, uno en C, uno en G, cuatro en I, dos en L, dos en O, uno en Q1 y dos en Q3. Se elaboró la plantilla y se aplicó a las muestras de ambos sexos y situaciones del estudio de distorsión: 465 protocolos y a otros 612 sujetos de procesos de orientación y selección. Se ha determinado el punto crítico $X=6$ con el cual parece que comienza la tendencia a la variable implicada.

Con estas distribuciones, se han elaborado los intervalos del siguiente cuadro.

Punt. en Escala Negación o Azar	Interpretación tentativa	Porcentaje incluido
0 - 5	Cooperación	93.5 %
6 - 9	negación	5.66 %
10 - 22	negación alta	0.84 %

Quando en los sujetos "sinceros" en DM se separa el grupo de "cooperación" sólo queda el 1.21% y no hay ninguno con "negación" alta. En la muestra de "distorsionadores" este baremo discrimina un 2.31% con "negación" y solo un 0.91% con "negación" alta. (Seisdedos, 1985) Cuando esta escala de negación se correlaciona con las propias del 16 PF no se observan relaciones significativas; esto parece comprensible si se tiene en cuenta que la denegación (aunque toma elementos de dichas escalas) se apoya en las alternativas más infrecuentes. (Seisdedos, 1985) p 42,43

ESCALA DE "MALA IMAGEN" o DISTORSIÓN NEGATIVA DEL 16 PF-A.

Otro tipo de distorsión deliberada es la tendencia de las personas por aparecer tan mal como les es posible, por ejemplo, en los casos en que una persona busca psicoterapia o se ven forzadas a evitar una situación desagradable. Este tipo de distorsión es poco frecuente en la evaluación de personal pero aparece en la evaluación con fines legales y educativos, entre otros. A esto se le llama Distorsión Negativa. (Karson y O'Dell, 1989) En los experimentos para estudiar esta escala, las instrucciones de falseamiento demandan que los sujetos respondan intentando presentar una mala imagen de ellos mismos. El procedimiento para construir esta escala es el mismo del de la escala de DM, la única diferencia la marcan las instrucciones.

La referida investigación de Winder et al (1975) derivó una escala DN que incluye los siguientes elementos. Se menciona el número consecutivo, el número del reactivo en el cuestionario, la alternativa -en minúscula- y la escala del 16 PF a la que pertenece -en mayúscula.

ITEMES QUE FORMAN LA ESCALA DISTORSIÓN NEGATIVA DN Winder et al, (1975)

1	14	a	M	9	89	c	L
2	38	a	L	10	117	a	N
3	42	c	N	11	119	a	O
4	51	b	A	12	123	a	Q3
5	52	c	A	13	143	a	O
6	55	c	C	14	176	c	A
7	68	c	O	15	182	c	F
8	80	a	C				

Winder et al (1975) encontraron que solamente el 1% de 283 perfiles contestados en condiciones estándar, pudiera ser clasificado como falseado en la dirección negativa contra 7% de protocolos "espontáneamente" falseados en dirección positiva en la misma muestra, considerando un punto de corte de 6. De acuerdo con el análisis factorial en el mismo estudio, la escala de DN su ubica sola en su propio factor de segundo orden, el número VI, y corresponde por sí sola a una dimensión bastante diferente del falseamiento en una dirección positiva.

En seguida se muestran los resultados encontrados por Krug (1978) en EE.UU al aplicar dicha escala de distorsión negativa DN en una muestra de 4830 sujetos obtenida al azar de los servicios de evaluación rutinaria del I.P.A.T. (Institute for personality & ability testing) que comprenden todo EE.UU en un año. Esta muestra es razonablemente representativa de todos los tipos de sujetos y propósitos de evaluación de la población en la que el 16 PF se usa en la práctica. El coeficiente de confiabilidad de la escala DN fue de $\alpha = .43$.

BAREMOS DE LA ESCALA DE DISTORSION NEGATIVA DEL 16 PF (KRUG, 1978)

puntaje bruto	puntaje esten	Rango potencial de respuestas: 0 a 15
0	1	Rango real en mujeres: 0 a 11 Media de 2.24 d s 2.04
1	5	Rango real en hombres: 0 a 12 Media de 2.45 d s 2.27
2-3	6	
4	7	
5-6	8	
7	9	
8-15	10	

La distribución resultó muy asimétrica y menos del 10% de los examinados obtuvieron puntajes superiores a 6 puntos. Dado que no se encontraron diferencias significativas entre los puntajes de ambos sexos, la tabla de baremos para convertir el puntaje bruto de DN a puntuación esten se obtuvo reuniendo los puntajes de ambos sexos.

Krug (1978) propone una forma de corrección de los puntajes de algunas escalas del 16PF que son afectadas por la DN.

- Para un puntaje esten de 7, se agrega un esten a la escala C
- Para un puntaje esten de 8, agregar un esten a C y restar uno a las escalas O y Q4
- Para un puntaje esten de 9, agregar un esten a C, H, I y Q3. Restar uno a L, O y Q4.
- Para un puntaje de 10, se agregan dos estenes a C, un esten a A, H, I y Q3. Restar uno a L, O y Q3

En este capítulo se ha revisado el aspecto teórico de la distorsión de las respuestas a los tests de personalidad, los diversos tipos de distorsión, las formas de identificarla y los métodos para evitarla o reducirla. Se han explicado las escalas de validez como una alternativa para identificar la distorsión en los inventarios, tanto la direccionalidad como la magnitud. Se han expuesto las escalas de validez del MMPI y las del 16 PF para finalizar con la escala de DM. Asimismo, se han reseñado algunas investigaciones cuyos hallazgos se consideraron importantes para el propósito de este trabajo dentro de las cuales se refirieron algunos estudios españoles que adquieren especial importancia para México por razones idiomáticas y de similitud cultural. Desafortunadamente se encontró muy poca información de estudios recientes realizados en México sobre el tema.

Las escalas de DM de Azar y las de Distorsión Negativa aquí mencionadas son alternativas sencillas y prácticas para evaluar cierto tipo de distorsión en el 16 PF, no así los procedimientos propuestos por Catell y sus colaboradores. La fórmula de sabotaje y la ecuación de distorsión pues son procedimientos de difícil aplicabilidad.

3.3.4 LAS ESCALAS DE VALIDEZ DEL 16 PF-5

En 1993 se publica la nueva versión de la más importante prueba de Catell, el 16 PF-5 (llamado así, según el manual, por ser la 5ª. versión) circunscrito dentro del enfoque de los "CINCO GRANDES" pues considera cinco factores de segundo orden, en lugar de los cuatro de la versión anterior. Aquí ya se incluyen tres escalas de validez encaminadas a evaluar algunos estilos de respuesta como los "enfoques actitudinales ante el instrumento", las respuestas infrecuentes y la condescendencia. Seisdedos (1995).

Escala MI. Manipulación de imagen

La primera escala, llamada MI (del inglés Impression Management que en España tradujeron como Manipulación de Imagen para conservar las siglas originales) se formó con elementos seleccionados de anteriores escalas y con nuevos reactivos.

En la construcción de esta escala se intentó evaluar componentes conscientes e inconscientes del falseamiento. Una puntuación alta indica una mayor tendencia a responder de un modo sociable o una exageración de esas cualidades. Una puntuación baja, por el contrario, muestra la tendencia a adjudicarse aspectos socialmente menos deseables. Se formó con elementos existentes previamente y algunos nuevos. Los doce elementos finales de la escala sólo puntúan en esta escala y son independientes de las escalas de personalidad del 16 PF (Seisdedos, 1995) lo cual constituye una innovación metodológica ya que las anteriores escalas se formaban con los propios elementos del inventario.

La escala se validó al aplicarla conjuntamente con otras pruebas: a) el Inventario BIRD de Paulhus, cuyas subescalas evalúan supervaloración del yo y manipulación de imagen, b) La escala de Marlowe-Crowne que evalúa deseabilidad social y c) las escalas de buena imagen y de mala imagen del 16 PF. Se obtuvieron correlaciones entre .45 y .55. Se hizo un análisis factorial conjuntamente con los factores de 2o orden con la intención de identificar qué rasgo de personalidad es más característico de tal conducta y se encontró que tiene que ver con Ajuste, con Autocontrol y con Independencia en sentido inverso (Siesdedos, 1995) En un estudio de esta escala realizado por Conn y Rieke (1994, en Seisdedos, 1995) se encontraron diferencias significativas entre las aplicaciones en dos situaciones, una anónima y otra de selección de personal así como correlaciones con los rasgos primarios de C fuerza del yo y Q4. Por supuesto, esta escala se normalizó tanto en EEUU como en España por lo que se cuenta con los baremos correspondientes. Por cierto, la media de EEUU fue inferior a la encontrada en España.

Escala de infrecuencia

La segunda escala, llamada Infrecuencia IN se formó con los 18 elementos del test elegidos bajo un criterio de infrecuencia, es decir, que solamente el 4% o menos de la población haya presentado esas respuestas. Una puntuación alta indica que el sujeto contestó de forma bastante inusual y las probables explicaciones son: respuestas dadas al azar, una pobre comprensión de la lectura o un intento de evitar dar una mala interpretación. (Seisdedos, 1995)

La escala de Aquiescencia

La tercera escala, llamada Aquiescencia AQ (se prefiere traducirlo como condescendencia) mide la tendencia a contestar como verdadero a un número inusual de elementos, independientemente de su contenido. Se forma con 95 elementos del propio test. Evalúa la incongruencia al contestar -contestar afirmativo a elementos con contenido contradictorio- o una fuerte necesidad de aprobación del sujeto frente al examinador. (Seisdedos, 1995)

En el siguiente capítulo se planteará el aspecto metodológico de una de las escalas de validez del 16 PF la escala de DM ya que se describirá el procedimiento seguido por Winder, O'Dell y Karson para construir tal escala en 1975, cuya vigencia se puede constatar al encontrar investigaciones acerca de ella a lo largo de los años hasta el reciente 1994 (Christensen et al) También se referirán otras investigaciones acerca del tema.

CAPITULO 4.

EL METODO DE WINDER, KARSON Y O'DELL PARA DERIVAR UNA ESCALA DE DISTORSION MOTIVACIONAL PARA EL 16 PF-A.

En este capítulo se expone el método seguido por Winder, Karson y O'Dell para derivar escalas de validez del 16 PF. Aunque el propósito de este trabajo es identificar la distorsión en sentido positivo, llamada distorsión motivacional DM, de los sujetos que contestan a tal cuestionario en situaciones laborales, aquí se presenta el método seguido por los autores mencionados para construir dos escalas de distorsión, que identifican tanto el falseamiento positivo como el falseamiento negativo, que es básicamente el mismo.

4.1 Antecedentes

Ya se había comentado que Catell y sus colaboradores diseñaron escalas de D.M para las formas C y D del Cuestionario 16PF pero no lo hicieron para las formas A y B, que son las más usadas con fines de evaluación de personal. Winder, Karson y O'Dell se dieron a la tarea de construir escalas de distorsión para la forma A del citado 16 PF.

Estos autores (1975) refieren estudios sobre el falseamiento del 16 PF realizados por Braun y LeFaro (1968, 1969), Catell et al (1970), Meredith, (1968) y Karson (1967) pero sólo este último dio lugar a la construcción de una escala de falseamiento, la construida por Schanberger y Ciotola (citados por Karson, 1967, en Winder et al, 1975) para identificar específicamente la Distorsión Motivacional. Es este el antecedente instrumental del que parten Winder y sus colaboradores para construir dos escalas de falseamiento, una para detectar el falseamiento positivo, cuando el sujeto desea dar una mejor impresión de sí mismo, llamada Distorsión Motivacional; y la otra para detectar el falseamiento negativo, en los casos en que el respondiente desea dar una peor impresión de sí mismo, llamada por los autores Distorsión Negativa, DN. Esta última no debe confundirse con la Escala de Negación estudiada por Seisdedos (1985), pues este nombre se refiere más bien a la situación de "negarse a colaborar" y la escala se forma con las alternativas atípicas de respuesta.

Las dos escalas de distorsión referidas se derivan de los reactivos del propio cuestionario 16 PF forma A. Krug (1978), se refiere a ellas como escalas autónomas y empíricamente validadas.

A diferencia de la escala DM de Schanberger y Ciotola, que fue derivada de las respuestas de psicópatas hospitalizados, Winder y sus colaboradores (1975), lo hicieron a partir de las respuestas de personas normales, en este caso, estudiantes universitarios. El mismo procedimiento lo aplicaron a la construcción de la escala de distorsión negativa.

El procedimiento consiste, de manera resumida, en identificar los ítemes que mejor discriminan entre tres grupos, uno que contestó el inventario bajo las instrucciones usuales de sinceridad, y los otros dos que lo contestaron bajo instrucciones de falseamiento, ya fuera para presentar una imagen mejorada o empeorada de sí mismo. De los protocolos contestados en las dos situaciones de falseamiento, se identifican dos conjuntos de quince reactivos que forman las escalas; al calificar éstas en los protocolos del 16 PF, se obtienen los puntajes de distorsión en ambas direcciones. Adicionalmente se obtienen los puntajes de distorsión en los protocolos de otra muestra de estudiantes, que contestó el inventario tanto en condiciones de sinceridad como en condiciones de falseamiento en ambas direcciones, para verificar su poder de discriminación y obtener una validación cruzada. Se cuenta asimismo con una muestra mayor de personas que contestaron en condiciones de sinceridad y que sirve de grupo de comparación.

En una muestra mayor de protocolos de 16 PF ya contestados, se califican las nuevas escalas, se factorizan los puntajes brutos de las 16 escalas primarias y las dos nuevas escalas y se obtienen las razones F y el porcentaje de varianzas con el fin de identificar las escalas que son afectadas por las tres condiciones de prueba: la condición usual de sinceridad, la de distorsión motivacional DM , y la de Distorsión Negativa DN y si es significativa es esa afectación Winder y sus colaboradores concluyen proponiendo un puntaje de 6 o superior, en las mencionadas escalas, como indicador de distorsión e identificando las escalas primarias que son más afectadas por la distorsión.

4.1.1 DESCRIPCIÓN DEL PROCEDIMIENTO.

A continuación se describe paso a paso el procedimiento para la construcción de tales escalas.

Para la identificación de los ítemes que forman las escalas de distorsión, una primera muestra $n=82$ de estudiantes universitarios se dividió en dos grupos. Al primero, $n=45$, se le aplicó dos veces el 16 PF-A en una sola sesión. La primera aplicación fue con instrucciones de contestarlo como si fueran a ser elegidos para el alistamiento militar o para obtener una compensación por incapacidad (condición de distorsión negativa). La segunda aplicación fue con instrucciones de contestarlo como si fueran candidatos a un empleo que les interesara bastante (distorsión positiva). En ambos casos se les dijo que había formas de detectar la distorsión, de tal manera que ésta debería ser fina y sutil. Esta es una situación de falseamiento simulado, inducido, no real. Para elegir los quince reactivos de cada escala, se siguieron los criterios de Shanberger y Ciotola a saber: a) Más del 50% de los sujetos deben haber contestado en esa dirección en la condición de falseamiento. b) Debía haber más del doble de respuestas en esa dirección en la condición de falseamiento que en la condición estándar.

Se contaba con una segunda muestra de protocolos del 16 PF-A de 283 estudiantes universitarios que servían como grupo estándar de comparación. De este primer grupo de 45 estudiantes se obtuvieron las escalas de distorsión iniciales. A continuación se muestra la lista de reactivo que forma la escala, que corresponden al 16 PF forma A que actualmente se usa en México.

Se menciona el número consecutivo, el número del reactivo en la prueba, la alternativa y la escala primaria a la que pertenece.

REACTIVOS QUE FORMAN LA ESCALA DE DM DE WINDER ET AL (1975)

1.	7c	E	9.	123c	Q3
2.	24c	Q3	10	130a	C
3.	61c	H	11	133c	F
4.	62a	I	12	149c	Q4
5.	81a	E	13	173a	Q3
6.	97a	Q2	14	174c	Q4
7.	111a	H	15.	184a	G
8.	114c	L			

Winder et al (1975)

Procedimiento para la validación cruzada de las escalas

Estas escalas fueron revalidadas en el resto de la primera muestra, $n=37$. Esta fue subdividida en dos grupos; al primero de ellos, $n=20$, se le aplicó dos veces el cuestionario: la primera aplicación con instrucciones de falseamiento positivo y la segunda con las instrucciones usuales de honestidad (se les garantizaba que sus respuestas permanecerían anónimas)

Al segundo grupo también se le aplicó dos veces el 16 PF-A, la primera aplicación con instrucciones de falsear negativamente y la segunda con instrucciones de responder honestamente. En todas las aplicaciones, se revirtió el orden en aproximadamente la mitad de los sujetos para contrabalancear cualquier efecto del orden de aplicación. Estos dos últimos grupos sirvieron para una validación cruzada de las escalas iniciales.

Resultados. Se seleccionaron los dos grupos de 15 ítems que mostraron la mayor variación con el falseamiento, tanto positivo como negativo. Se aplica la fórmula de t para calcular si las diferencias encontradas entre las condiciones de falseamiento y las condiciones de aplicación estándar son significativas con una probabilidad $t < .001$

En este estudio, se encontró una diferencia significativa entre las medias de las escalas de DM en la situación de falseamiento y en la situación estándar. También hubo diferencias significativas en las medias de la escala de DN en la situación de falseamiento y en la situación estándar. Estos autores proponen que un puntaje bruto de seis puntos o mayor en cualquiera de las dos escalas puede ser considerado como indicador de falseamiento.

El Análisis factorial de las escalas

Una vez que las dos escalas de falseamiento han sido desarrolladas y se ha realizado su validación cruzada, se realizó un análisis factorial de estas junto con las escalas originales del 16 PF en una nueva muestra de 245 estudiantes universitarios. La factorización fue hecha como sigue: se calcularon las correlaciones entre las 18 escalas usando sus puntajes brutos y se insertaron las correlaciones múltiples en la diagonal para estimar las comunalidades. Estas fueron además refinadas a través de cinco iteraciones de comunalidad. Se hizo la rotación hasta la estructura simple a través de OBLIMIN.

Se rotaron seis factores aprox. siguiendo la sugerencia de Veldman (1967 en Winder et al 1975) de que seis es el más justificable número de factores de segundo orden que puedan encontrarse en el 16 PF.

Se realizó un análisis de varianza de cada escala y se obtuvieron las razones F para determinar la significancia de las diferencias entre las varianzas en las escalas originales del 16 PF y las nuevas escalas de distorsión. Los autores sin embargo, consideran que es necesario obtener la medida de la intensidad del efecto en las escalas debido a que éstas serán significativas por el tamaño de las muestras y proceden a obtener la proporción de la varianza atribuible a las diferencias. Decidieron arbitrariamente que si una comparación no explica al menos el 10% de la varianza, sería considerada como de poca importancia para fines prácticos.

La afectación de distorsión en las escalas primarias.

Usando este estándar del 10 % de la varianza, se encontró que las escalas E, I, M, N, Q1 Y Q2 del 16 PF son más inmunes a la distorsión en los grupos de estudiantes universitarios ya sea en la dirección positiva o en la negativa. Las escalas más susceptibles al falseamiento, además de las nuevas escalas de distorsión, son C, G, L, O, Q3 y Q4. De suficiente interés es el notar, exceptuando G, que todas ellas cargan (saturan) fuertemente en el factor de segundo orden de ansiedad (Catell y otros, 1970). Así, de acuerdo con esta investigación, se encontró que algunas escalas del 16 PF están comparativamente libres de falseamiento y otras están fuertemente afectadas por este. Es obvio que la nueva escala de DM, se vincula con el factor de segundo orden Integración Dinámica vs Ansiedad. Esto implica que la Distorsión Motivacional está cercanamente vinculada con la fuerza del ego y con la negación de la ansiedad.

La estructura factorial de las nuevas escalas, factorizadas junto con las escalas originales del 16 PF, muestran claramente que los resultados siguen el patrón usual de los factores de segundo orden del 16 PF. Así, el primer factor es fácilmente reconocido como Integración dinámica vs. ansiedad, con cargas altas en C-, L+, O+, Q3-, Q4+ y MD-.

El segundo factor es bien conocido en la literatura del 16 PF como Sometimiento vs Independencia, con cargas en E+, L+, N- y Q1+ (La inclusión de N aquí es atípica) El tercer factor es Exvia vs Invia, con cargas en A, F H y Q2. Los factores restantes siguieron los patrones usuales del 16 PF con excepción de un factor de segundo orden diferente, que satura la escala de DN. Se identificó una sola dimensión bastante distinta de la DM y esto implica que esta escala de distorsión negativa mide algo único, diferente y no sólo un tipo de distorsión contraria a la DM.

Aquí es conveniente señalar que, en una situación de aplicación normal, se espera un cierto porcentaje de sujetos "espontáneamente" falseadores. En la muestra de 283 perfiles contestados en condiciones estándar y considerando el puntaje de 6 como indicador de distorsión, el 7% de la población normal debe ser considerada como falseadora en la dirección de DM, y esto puede considerarse como un indicador adicional de la eficacia de dicha escala. En opinión de Seisdedos (1985) el índice de distorsión debe obtenerse de la diferencia entre la diferencia entre los puntajes de una situación de distorsión y los puntajes de una situación normal, y no solo de la primera, como es lo usual. Pero en una situación motivante o dudosa, rara vez se cuenta con ambos valores y la alternativa es tomar como referencia baremos de distorsión.

Winder y cols. recomiendan que estas escalas deberían ser usadas para una evaluación preliminar de la distorsión pero agregan que debe tenerse cuidado al hacer generalizaciones a otros grupos debido a que estas escalas fueron derivadas de respuestas de estudiantes universitarios. Una evaluación completa del falseamiento requiere que se considere tanto la personalidad como las variables situacionales. Los evaluadores no deberían caer en la trampa de creer que una persona está falseando sus respuestas simplemente por un alto puntaje en la escala de DM o de DN.

Siempre que sea posible debe considerarse cada caso particularmente y apoyarse en otros instrumentos o medios para decidir si la distorsión llega a invalidar el perfil. El estudio de la situación en que se aplica el test es también un elemento que ayuda a analizar la distorsión y a considerar su impacto en los resultados de la evaluación psicológica completa.

4.2 EL ENFOQUE CORRECTOR DE KRUG

¿Qué se hace con los protocolos que obtienen un puntaje alto en distorsión? ¿Es permisible eliminar perfiles por el solo hecho de que sus puntajes de distorsión sean altos? Una alternativa de solución la presenta Krug (1978) al proponer un sistema de corrección de las escalas primarias dependiendo del puntaje de las escalas de distorsión de Winder y cols..

Samuel E Krug es un miembro destacado del Instituto de Evaluación de la Personalidad y la Habilidad, IPAT, el instituto creado por Catell Krug (1978) califico las escalas de distorsión de Winder et al en una muestra bastante mayor ($n=4830$) y más representativa de la población a la que usualmente se aplica el 16 PF que la muestra de estudiantes universitarios estudiada por Winder y cols. En esta muestra estaban representados diferentes propósitos de evaluación, no había casos de distorsión simulada aunque puede suponerse que sí había casos de distorsión motivacional pues algunas evaluaciones perseguían fines laborales. En este estudio, Krug obtuvo puntajes diferenciales por sexo, baremos de distorsión positiva y negativa, un índice de confiabilidad de las escalas y presentó una propuesta de corrección de las escalas primarias por los efectos de la distorsión.

Krug (1978) seleccionó una muestra al azar de los protocolos del 16 PF-A reunidos en los servicios de evaluación de IPAT durante un año. Obtuvo puntajes de las dos escalas de distorsión y los analizó diferencialmente por sexos. Aunque obtuvo una t significativa entre los puntajes de hombres y mujeres, quizá por el tamaño de las muestras, una prueba adicional: w_2 (Hays, 1963 en Krug, 1978) no indico diferencias significativas entre sexos por lo que reunió los puntajes y obtuvo baremos en estenes, que se muestran en el siguiente cuadro.

BAREMO PARA LA DISTORSION MOTIVACIONAL DE KRUG

puntaje bruto DM	estén	
0	1	Hombres media 6.36 d s 2.87
1	2	
2-3	3	Mujeres media 5.71 d s 2.27
4	4	
5	5	
6-7	6	
8-9	7	
10	8	
11	9	
12-15	10	

Krug (1978)

A partir de sus hallazgos, Krug (1978), consideró que el puntaje bruto de 6 o más que Winder et al (1975) habían considerado como indicador de distorsión en ambas direcciones era demasiado bajo en la escala DM pero aproximadamente correcto en la escala DN. Si se utilizara tal criterio, aproximadamente el 55% de todos los individuos rutinariamente examinados con la prueba sería juzgado como falseador en dirección positiva, en lugar del 7% reportado por Winder. Krug propone un puntaje bruto de 10 en la escala de DM como punto de corte "si fuera necesario contar con uno".

Adicionalmente, él encontró que los índices de confiabilidad de ambas escalas indicaban la suficiente consistencia como para permitir que se hicieran ajustes en los rasgos primarios cuando la distorsión fuera superior al promedio.

Para determinar el impacto que las dos formas de distorsión tienen sobre las escalas primarias del 16 PF y las correcciones que pudieran ser posibles, cada puntaje de falseamiento se correlacionó con las 16 escalas de rasgos primarios. Los patrones de correlación encontrados son congruentes con lo esperado: en términos generales, la gente que intenta presentarse bajo una luz más favorable se muestra menos ansiosa; específicamente, aparece como más estable emocionalmente, más audaz, menos sospechosa y menos culpable, más autocontrolada y menos tensa. Es decir, traducido al lenguaje factorial del 16 PF: C+, H+, L-, O-, Q3+, y Q4-.

Las correlaciones de las escalas de falseamiento con los factores primarios del 16 PF resultarían muy congruentes con las obtenidas en el estudio de Winder et al (1975), y proveen una evidencia adicional de la validez de ambas escalas.

Krug (1978) sugiere los siguientes ajustes basados en las regresiones de las escalas de distorsión con las escalas de los rasgos primarios. Estos ajustes representan un intento de separar la distorsión de la varianza real de los rasgos. Estos ajustes o puntajes de corrección se deben usar con precaución.

CORRECCION DE PUNTAJES DE LAS ESCALAS PRIMARIAS DEL 16 PF-A

DM = 7, agregar un 1 a C4 y sustraer 1 de C

DM = 8, agregar 1 a L, O, y Q4, sustraer 1 de A, C, G, y Q3

DM = 9, agregar 1 a F, L, O, y Q4, sustraer 1 de A, C, G, y Q3

DM = 10, agregar 2 a Q4, sustraer 2 de C, agregar 1 a F, L, y O, sustraer 1 de A, G, H, y Q3

No se indican correcciones cuando el estén de DM es menor a 7.

Krug (1978)

La propuesta de corrección de Krug adquiere una mayor importancia en los casos de evaluaciones grupales en que el tiempo es un elemento escaso pues permite un manejo cuantitativo de la distorsión y el psicólogo puede tener una base consistente -uniforme para todos los evaluados- para tomar decisiones.

Por otro lado, Kline (1993b) señala que, en un proceso de selección de personal, todo lo que se puede hacer con individuos que distorsionan deliberadamente sus respuestas en un inventario es ignorar sus puntajes en los campos en que probablemente éstos sean inexactos. Concluye que no es posible corregir los puntajes por efectos de la distorsión.

4.2.1 UNA APLICACION DE LA CORRECCION DE KRUG EN ESCENARIOS LABORALES.

Este enfoque de corregir los puntajes de las escalas del 16 PF-A a partir de los puntajes de distorsión sí puede tener consecuencias importantes en escenarios aplicados como la selección de personal, como lo demuestra un estudio de Christiansen y Cols (1994) que también utilizaron las escalas de distorsión de Winder et al (1975) y el sistema de corrección de Krug (1978). Estos investigadores se propusieron estudiar el efecto de los puntajes corregidos por un lado, en una validación con relación a criterio y por otro, en las decisiones de contratación de personal. Aunque normalmente se asume que las correcciones por DM tienen un efecto benéfico en la validez, no existe una evidencia de investigación que apoye esta suposición. Conceptualmente, el uso de puntajes de corrección ubican a las escalas de distorsión como variables supresoras que eliminan la varianza de los rasgos que se suponen como irrelevantes para descubrir la relación entre los predictores y el criterio (Christensen, et al. 1994).

La investigación de Christensen et al (1994) consistió en lo siguiente: aplicaron 16 PF-A en un contexto organizacional a un grupo de empleados en funciones y candidatos externos a puestos altos de supervisor, en total 495 personas. Se contaba con la evaluación de desempeño de 84 de los empleados la cual fue el criterio considerado para el estudio de validez de los rasgos de personalidad relevantes al trabajo, identificados a partir de un análisis de puesto. Se consideraron cinco factores de segundo orden del 16 PF, calificados con y sin correcciones por distorsión. Los resultados indican que la corrección del falseamiento tuvo un pequeño efecto en la validez con relación a criterio pero no fue importante: no hubo un aumento significativo de la varianza del criterio explicada por los puntajes corregidos que la explicada por los puntajes no corregidos.

Aunque las decisiones reales de selección se tomaron con base en los resultados de simulacros assessment-center, los efectos de la corrección de las escalas del 16 PF en las decisiones de selección pudieron examinarse por situaciones hipotéticas considerando diferentes razones de selección. Dependiendo de estas últimas, las decisiones de contratación (hipotéticas) individuales sí fueron afectadas ya que los puntajes corregidos por falseamiento dieron lugar a diferentes resultados de los que hubieran dado los puntajes no corregidos, es decir, un individuo que hubiera sido seleccionado sobre la base de los puntajes del 16 PF sin corrección por falseamiento, no hubiera sido seleccionado con base en los puntajes corregidos por falseamiento.

A pesar de que la investigación inicial especialmente enfocada a la distorsión concluyó que ésta no afecta típicamente las respuestas a las medidas de personalidad. (Hough et al, 1990, en Christensen et al, 1994). Se observa claramente la necesidad de realizar investigación adicional para determinar si la distorsión representa una seria amenaza a los tests de personalidad y si se justifica la corrección por distorsión. Si la DM probara ser rara o infrecuente, resultaría innecesario el uso rutinario de escalas de distorsión para "corregir" las respuestas.

Existen razones para dudar que sea benéfico el uso de la corrección de los puntajes cuando se aplica rutinariamente en la evaluación en contextos organizacionales. Las escalas de validez suponen que existe un solo conjunto de respuestas de distorsión motivacional DM en todos los puestos pero los rasgos para predecir el desempeño varían de puesto a puesto y los individuos tienen diferentes ideas de qué rasgos serán percibidos como favorables según el puesto; esto da como resultado diferentes patrones de distorsión. Además, los puntajes de corrección generalmente han surgido para ayudar en la evaluación individual y los efectos al aplicar estas correcciones en grupos, en un contexto de selección de personal, no son suficientemente claros. (Christensen et al, 1994)

Efectivamente, el hecho de tomar en cuenta puntajes corregidos o no, puede tener un impacto importante en las decisiones organizacionales que se basan en los puntajes de un test de personalidad. Pero este y otros riesgos mayores existen al considerar para las decisiones organizacionales únicamente el resultado numérico de un inventario aunque se cuente con el respaldo metodológico y del tratamiento estadístico de los puntajes tan completo como el utilizado por estos investigadores.

4.3 DERIVACION DE UNA ESCALA DE D.M. EN ESPAÑA.

Con el procedimiento seguido por Winder y cols (1975), Seisdedos (1985) informa de un estudio de construcción de una escala de distorsión motivacional DM realizado en España

Se aplicó el cuestionario 16 PF a una muestra de 57 varones y 157 mujeres en dos ocasiones. En la primera ocasión se solicitó su sinceridad. En la segunda se les pidió que fingieran ser candidatos a un empleo que les interesara mucho. Se analizaron las respuestas y se identificaron aquellas quince cuestiones que más se ven afectadas por dicho fenómeno de distorsión, es decir, que discriminaran mejor las dos situaciones.

REACTIVOS DE LA ESCALA DE DISTORSION MOTIVACIONAL DM PARA EL 16 PF-A (Seisdedos, 1985)

4a	35c	75a	145c
10a	61c	97a	174c
24c	62a	99c	180a
33a	73a	147c	

Nota: Los números de los reactivos corresponden a la versión en inglés y a la de México.

En las siguientes figuras se muestra gráficamente el impacto de la distorsión en las escalas primarias del 16 PF en ambos sexos.

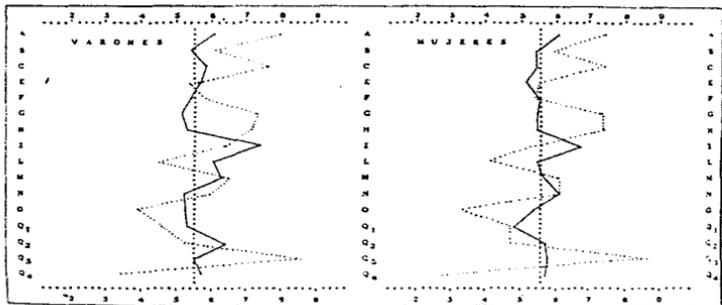


Figura 2.—Perfil sincero (—) y distorsionado (-----) en 15 PFA.

Sesdedos (1985) p 35

En esta figura están representados en decatipos los perfiles obtenidos por ambos sexos en las dos situaciones experimentales de aplicación: bajo instrucciones de sinceridad y de distorsión. La línea vertical de los puntos alude a los datos normativos de la tipificación española. Puede verse gráficamente cómo los sujetos distorsionadores elevan las escalas A+, C+, G+, H+, y Q1+ y en el otro polo, destacan la L-, O- y Q4-.

Sesdedos (1985) agrega que se han analizado también las distribuciones de la Escala de distorsión motivacional DM en todas las muestras estudiadas, cuyo número es superior a 1200 y tentativamente se ha definido como punto crítico una puntuación DM=7. Con ella se discriminan aproximadamente cuatro de cinco sujetos distorsionadores, aunque también queda incluido uno de cada cinco sujetos sinceros. Dichas diferenciaciones han permitido definir los intervalos siguientes.

BAREMO DE DISTORSIÓN MOTIVACIONAL Sesdedos (1985)

- 0-3 Muy sincero
- 4-6 Sincero
- 7-10 Algo distorsionador
- 11-15 Distorsionador

La calificación de la escala DM en los grupos experimentales arrojó los siguientes resultados:

ESCALA DM EN SUJETOS SINCEROS Y DISTORSIONADORES (n=214)

	SINCERIDAD		DISTORSION	
	varones	mujeres	varones	mujeres
media	5.53	3.56	10.93	9.99
d.s	2.87	2.50	3.96	4.23

(Seisdedos, 1985)

Esta escala DM se revalidó en una segunda muestra de 815 sujetos de ambos sexos que contestaron el 16 PF en situaciones reales de selección y de orientación. Esta última muestra es diferente en cuanto a que la situación de falseamiento no es inducida artificialmente sino real por el hecho de que los protocolos fueron obtenidos de aplicaciones con propósitos de selección o de orientación.

RESULTADOS DE DM EN DOS SITUACIONES (n=815)

	ORIENTACION		SELECCION	
	Varones	Mujeres	Varones	Mujeres
N	199	195	380	41
media	6.17	4.72	8.69	8.02
d.s	2.93	2.9	3.39	3.42

(Seisdedos, 1985)

También hubo diferencias entre los dos grupos que contestaron 16 PF con diferentes propósitos. El de orientación resultó sincero en promedio y los de selección entraron en el grupo "algo distorsionador". Como en la situación de sinceridad de la muestra experimental, se observa que las mujeres obtuvieron puntajes menores en DM que los hombres.

Además, se realizó un análisis correlacional entre los puntajes de las 16 escalas originales del 16 PF en los grupos de sujetos sinceros y distorsionadores. Los resultados, como lo refiere Seisdedos(1985) concuerdan con las afirmaciones de Catell de que la gente intenta ser considerada con valores altos en A, en C, en G, en H y en Q1 y por otro lado, la gente desea mostrar valores bajos en L, O, Q4 y, menos claramente en el estudio español, en Q2 y M. Esto se traduce, en términos de factores de segundo orden, en una evitación de la Ansiedad y una tendencia hacia la Extroversión exceptuando F.

El procedimiento de Winder et al (1975) para desarrollar escalas de falseamiento tiene varias ventajas que se mencionan a continuación. La primera es que estas escalas se forman con una selección de los *mismos* reactivos del inventario que implícitamente aprovechan su valor estadístico en cuanto a confiabilidad, validez, objetividad, etc. lo cual conduce a resultados cuantitativos y susceptibles de ser convertidos a estenes. La segunda ventaja estriba en que el procedimiento para identificar los ítemes es relativamente fácil de realizar, permite una validación cruzada y el cálculo de la confiabilidad. La tercera ventaja radica en el hecho de que las escalas así derivadas son susceptibles de tratamientos estadísticos posteriores tan sencillos o tan complejos como lo permitan los recursos del investigador, desde una correlación simple hasta un análisis factorial, entre otros. Al revisar el estudio de Winder y sus colaboradores, no se encontró evidencia alguna que impida construir una escala DM para la forma B.

Asimismo se reafirma la necesidad de derivar escalas de validez propias para cada país pues entre las escalas derivadas por Winder(1971) y por Seisdedos (1985) siguiendo el mismo procedimiento, solamente coincidieron cinco de los 15 elementos que las forman.

CONCLUSIONES

1. LA EVALUACION PSICOLOGICA DE PERSONAL EN MEXICO.

La evaluación de personal en México se aplica a varios fines: la selección de personal, la promoción, la capacitación, el plan de carrera, entre otros. El primero es el más frecuente.

Los principales criterios que actualmente se utilizan para elegir un procedimiento selección son: la efectividad de los resultados, el costo y el tiempo. En muchas organizaciones mexicanas, el orden se revierte y por razones de costo y de tiempo, se eligen las técnicas de mayor error probable como son la entrevista abierta y la observación asistemática.

La efectividad del procedimiento de selección rara vez se hace de una manera formal y sistemática. Con frecuencia se toma el índice de rotación para valorar la selección pero la separación del empleo obedece a causas múltiples, no sólo refleja errores de selección del personal. Esta efectividad se estima de una manera manifiesta por los falsos positivos: el contratar personas que no tengan un desempeño adecuado de acuerdo con ciertos criterios, pero también debería evaluarse por los falsos negativos: el rechazar candidatos adecuados; pero este error es encubierto pues no se puede verificar. Y ambos son errores de igual magnitud.

Sin embargo, existen procedimientos de evaluación como la entrevista estructurada, la forma de solicitud ponderada y las pruebas objetivas, tanto de personalidad como de capacidades, entre otros, que resultan efectivos y no son ni costosos, ni tardados ni complejos.

Una proporción importante de organizaciones laborales utilizan algún procedimiento de evaluación psicológica de personal en México pero la pobreza de éstos se evidencia en una encuesta realizada en 1994 (Milán en Arias Galicia, 1996) pues los instrumentos más utilizados para evaluar la personalidad en la muestra de empresas estudiada son: Cleaver, 16 PF y las técnicas proyectivas de Machover y HTP. Otros instrumentos que se acostumbra utilizar son el MMPI, el Test de Luscher, Frases incompletas, el Moss, entre otros.

En mi práctica profesional he encontrado que es frecuente el uso deficiente de éstos y otros instrumentos, que los procedimientos de evaluación psicológica no son resultado de un análisis de las necesidades de las organizaciones y que los criterios para elegir pruebas no son los relativos a su rigor metodológico o a sus propiedades psicométricas. Asimismo, no se presta especial cuidado al proceso de evaluación en sí mismo, por lo que el error aleatorio potencial es muy alto.

Un problema muy extendido, que ya se había comentado, es el relativo al material de aplicación; por no utilizar el material original, se adicionan variables extrañas que afectan los resultados, usualmente en contra del evaluado.

Por supuesto, existen organizaciones que cuentan con sistemas muy avanzados de evaluación, que han adaptado pruebas, que utilizan técnicas multidimensionales como el simulacro assessment center y que además, constantemente están investigando y verificando el procedimiento, pero esas instituciones son más la excepción que la regla.

Problemas de la evaluación psicológica en México.

En México se carece de tecnología propia en materia de tests psicológicos, excepto en el campo de la evaluación de rendimiento en que se han construido instrumentos de amplia cobertura (ej. los desarrollados por la UNAM y recientemente por el CENEVAL para ingreso a educación media superior)

Son muy pocas las pruebas desarrolladas aquí con fines de evaluación de personal. Lo usual es que se utilicen instrumentos importados sin "tropicalizar", es decir, con una deficiente adaptación a nuestro país en lo referente a sus propiedades psicométricas, al idioma, a nuestra idiosincracia y a nuestras circunstancias.

En el campo de los tests cognoscitivos si existen adaptaciones importantes pero no sucede lo mismo con pruebas de personalidad de las cuales difícilmente se realizan estudios de confiabilidad, de validez y las normalizaciones, si se hacen, son deficientes. Se utilizan pruebas de los años 40's y los 50's, y se interpretan con un reducido conocimiento de ellas. No se hacen estudios de tipo lingüístico, imprescindibles en el caso de pruebas verbales diseñadas en otro idioma. Se usan asimismo, técnicas proyectivas importadas, diseñadas para evaluación clínica que requieren de "saltos mortales" para vincular sus resultados con necesidades de información organizacional. Se carece de datos acerca de la aplicabilidad intercultural de tales tests.

En los casos en que se realizan estudios serios de algún test, se cae en el gravísimo error de no difundir los resultados.

Las principales razones de este sombrío panorama, en mi opinión, son:

- a) la falta de preparación del psicólogo dedicado a la evaluación**
- b) la falta de comunicación entre organizaciones laborales y universidades**
- c) la falta de unión entre los psicólogos dedicados a este campo.**

La falta de recursos y de apoyo de parte de los empresarios no deberían ser un obstáculo sino un problema que los psicólogos preparados pueden superar con ingenio y creatividad. Muchos directivos son sensibles a una buena idea, bien fundamentada y práctica, que redunde en beneficio de la organización.

Se insiste en que el psicólogo dedicado a la evaluación debe formarse en la teoría, en la práctica y en metodología de investigación para que sobre la marcha realice investigación en su trabajo. La computación facilita sustancialmente esa investigación por lo que también debe contar con esta herramienta.

Una mayor unión y colaboración entre organizaciones laborales y universidades es imprescindible para conocer los avances mundiales en este campo, para enterarse de los estudios realizados en las universidades, para que los planes de estudio sean congruentes con las necesidades organizacionales, para realizar esfuerzos conjuntos de construcción de instrumentos propios, entre otros objetivos.

Y por último, el éxito de cualquier proyecto no es mérito de una sola persona. Los psicólogos deben unir sus esfuerzos para presentar un frente sólido ante la invasión de instrumentos extranjeros y para desarrollar tecnología propia aprovechando los avances de la Psicometría en todo mundo. Y, por qué no, formar equipos multidisciplinarios que persigan este fin y cuyos productos puedan ser aprovechados en otros países latinoamericanos.

Ética y comercialización

Otro aspecto importante de mencionar es el aspecto de ética y de comercialización, que en la realidad de la evaluación psicológica de México están estrechamente vinculados. En nuestro país no existe una reglamentación que norme la actividad de la evaluación psicológica, ni de parte de autoridades civiles, de universidades ni de asociaciones profesionales. Ni siquiera es requisito el contar con una cédula profesional de psicólogo para usar pruebas. Cualquier persona que se las arregle para conseguir tests puede usarlos y utilizar los resultados para diversos fines.

Al parecer, en México nos saltamos una etapa en el curso normal de existencia de los tests, que es la correspondiente a las consideraciones sociales y éticas de la evaluación psicológica y su consecuente reglamentación, que se habría dado naturalmente como en otros países, si no se hubiera presentado un fenómeno diferente propiciado por el tratado de libre comercio que es precisamente la comercialización. La ausencia de reglamentación en materia de evaluación psicológica convierte a México en un mercado potencial de enorme importancia para firmas extranjeras con desmedidos intereses económicos y restringidos escrúpulos de rigor científico. Estas firmas venden a cualquiera que pueda comprar y dejan instrumentos de dudoso valor en manos de personas sin preparación.

Todo parece indicar una relación inversa entre la agresividad de la comercialización y la seriedad del instrumento. Los tests serios se venden de una manera discreta a través de los distribuidores autorizados y restringen la venta a psicólogos acreditados. El costo es bastante razonable y accesible a cualquier empresa o psicólogo.

Estas firmas que últimamente han aparecido en nuestro país, encuentran un campo fértil para sus intenciones, abonado por la ignorancia de los profesionales de Recursos Humanos de los criterios elementales para elegir un test. En el medio laboral de las principales ciudades de México, circulan instrumentos que han sido vendidos aduciendo cualidades de rigor metodológico pero que en realidad son desconocidos en la literatura especializada. La promoción está a cargo de profesionales de la comercialización, no de profesionales de la psicología y el costo de esos instrumentos es usualmente bastante alto. Ya en 1984, Lanyon opinaba que los consumidores de este tipo de programas por lo general no son capaces de identificar un producto de calidad y que es preocupante que se haya convertido en norma la falta de validez demostrada de tales programas. (En Aiken, 1996)

El riesgo de tomar decisiones con base en los resultados de ese tipo de "pruebas" es incalculable si se considera la magnitud del error aleatorio de medida derivado de su construcción y falta de adaptación a México.

Lo anterior lleva al planteamiento de que existe una necesidad urgente de reglamentación de la evaluación psicológica, cualquiera que sea el fin que ésta persiga, así como de la necesidad de una certificación adicional del psicólogo que se dedique a tal actividad.

La práctica e investigación en Psicometría debe hacerse con una actitud socialmente responsable. El aspecto de la imparcialidad de las pruebas -oportunidad equitativa independiente de la raza, sexo, etc.- es medular. (Aiken, 1996) y aunque en México no existe reglamentación ni legal ni profesional en este sentido, es éste un aspecto que debe considerar cualquier organización o psicólogo profesional que tenga que ver con algún procedimiento de evaluación. El candidato a empleo es una persona que se ve sometida a procedimientos que con frecuencia son abiertamente parciales, injustos e inequitativos como el ser rechazado a priori por ser egresado de una universidad pública. Y no cuenta con recursos legales para defenderse de esta situación.

Los problemas de búsqueda de empleo masiva asociados con el altísimo nivel de desempleo de 1995 y 1996 desafortunadamente no dieron lugar a alguna innovación o reglamentación al respecto.

Los aspectos legales y técnicos que se derivan de una consideración de los conceptos de imparcialidad y predicción diferencial alertan a los psicólogos profesionales, a los gerentes de personal y al público en general con respecto de la necesidad de un uso más responsable de las pruebas. (Aiken, 1996) En EE.UU. y en otros países con tecnología psicométrica, las críticas y las leyes para regular e incluso prohibir el uso de los tests, han tenido el efecto de aumentar el cuidado en el diseño y la construcción de las mismas y han dado como resultado mejores pruebas.

2. LA EVALUACION MEDIANTE COMPUTADORAS.

Como uno de los factores que han contribuido al crecimiento -o al resurgimiento- de las pruebas psicológicas se ha mencionado al respaldo de sistemas computarizados en este proceso que no solo optimiza los resultados sino que abre posibilidades insospechadas en el tiempo y en el espacio. Aunque las computadoras se han utilizado en psicometría desde los años 50's, el advenimiento de las computadoras personales potenció su uso en este campo, igual que en muchos otros.

Sin embargo, una computadora no es muy útil sin una elección apropiada del software: el programa con el cual opera la computadora. La cantidad y diversidad de estos es enorme: tan sólo en EE.UU. se calcula que existen unos 200 programas computarizados para pruebas cognoscitivas, neuropsicológicas y de personalidad. (Aiken 1996)

Existen softwares generadores de pruebas para construir y calificar tests de aprovechamiento y de aptitud. La interpretación de los resultados de pruebas de personalidad es más compleja pero aún así, se han desarrollado programas para la aplicación, calificación, interpretación y generación de informes narrativos y gráficos. Los sistemas interactivos permiten una relación continua entre el sujeto y la máquina y son idóneos para la aplicación. La American Psychological Association APA, en EE.UU. ha emitido una serie de lineamientos aplicables a la administración de pruebas por computadora para que esta modalidad no sea desventajosa para algunas personas. (Aiken, 1996) En México aún no existe algún tipo de reglamentación relativa.

Los psicólogos que hacen uso de estos programas deben complementar los informes de los programas computarizados con afirmaciones interpretativas adicionales que se derivan de sus observaciones y experiencia, como lo recomienda la APA en EEUU. No es el propósito de estos programas el eliminar la intervención del psicólogo. Existe un daño potencial en los casos en que no esté presente un profesional que explique los resultados generados por los sistemas automatizados. Debe reconocerse además el hecho de que estos programas no siempre garantizan los requerimientos de rigor metodológico como la validez, la confiabilidad y la normalización. El problema aumenta debido a que la gente le confiere mucho valor a un reporte de resultados generado por computadora (Aiken, 1996)

Tales sistemas computarizados reducen sustancialmente las operaciones rutinarias del psicólogo, eliminan el error humano, permiten manejar un gran número de datos, reducen el tiempo de entrega de resultados al mínimo, van formando bases de datos que facilitan análisis estadísticos posteriores y otras ventajas adicionales, pero de ninguna manera pretenden excluir al psicólogo que tiene que estar, de principio, al frente del diseño o elección del sistema, pendiente de todo el proceso, que debe traducir los resultados generados por el software a los clientes internos en la organización y que finalmente debe ser el responsable del sistema y de sus resultados frente a la Dirección General.

Muchas pruebas están diseñadas para calificarse con facilidad y rapidez en forma manual, pero la computación proporciona un sinnúmero de ventajas adicionales. Sin embargo, no todas las pruebas son factibles de programarse para su operación computarizada, como ejemplo se mencionan algunas escalas del Weschler. Las computadoras también se utilizan en entrevista de psicodiagnóstico y en muchos campos más, facilitando de trabajo, ahorrando tiempo y optimizando algunos aspectos de la evaluación. (Aiken 1996) Cada programa debe evaluarse particularmente en función de las necesidades de la organización.

No son pocos los especialistas en evaluación que se muestran temerosos de los riesgos potenciales de la proliferación de estos programas y algunos de ellos han generado recomendaciones y sugerencias para evitar el ser sorprendidos por un vendedor de un producto que no cubre con los requisitos de rigor mínimos; asimismo alertan sobre un mal manejo de estos programas y recomiendan un énfasis en la capacitación de los usuarios. (Aiken, 1996)

La confidencialidad.

Uno de los problemas al utilizar estos programas es el garantizar la confidencialidad de los datos almacenados en la memoria de una computadora. En escenarios laborales, se van formando bases de datos sin la suficiente protección e incluso se almacenan en computadoras conectadas en redes, lo cual aumenta la probabilidad de un acceso no autorizado a los resultados individuales. (Aiken, 1996)

En el Reino Unido se emitió el Acta de Protección de Datos, referida por Kline (1993b) que establece el derecho de la persona a conocer sus resultados en caso de que éstos sean almacenados en una computadora, el derecho de los candidatos a recibir retroalimentación después del proceso de selección (los informes computarizados permiten hacer esto aunque sea un gran número de candidatos) y el garantizar a los evaluados la confidencialidad de sus datos. Al respecto, también se requiere una reglamentación en México.

Se recomienda a los usuarios potenciales de una prueba o programa que se informen de las características psicométricas del instrumento en cuestión y sobre todo, acerca de los datos de adaptación a México: qué estudios se hicieron, quién los hizo, cuando, dónde, qué tipo y tamaño de muestras, etc. Las obras de consulta mencionadas por Aiken, 1996, pp 9 a 21, son una excelente fuente de información para pruebas estadounidenses.

Futuro de la evaluación psicológica.

En todo el mundo, las pruebas psicológicas siguen prosperando y diversificándose, a pesar de -y gracias a- las críticas casi continuas, tanto internas como externas al ámbito de la Psicología. Las nuevas teorías, las nuevas pruebas, las revisiones de pruebas antiguas y los avances metodológicos que constantemente están apareciendo, dan testimonio del estado dinámico del campo. Como factores que han contribuido a este crecimiento tenemos el incremento de la población, la urgente demanda de aumentar la competitividad de las empresas y la consecuente necesidad de métodos más eficientes para evaluar a las personas en los contextos laborales, educativos y clínicos; la dinámica de la comunicación, entre otros. Esto también se ha visto estimulado por el progreso en la computación aplicada a la evaluación psicológica (Aiken, 1996)

En comparación con los avances técnicos de las pruebas de habilidades, los inventarios de personalidad y otros instrumentos afectivos son relativamente imperfectos. Sin embargo, es indiscutible el valor potencial de la evaluación precisa de la personalidad. El progreso reciente en los diseños con innovaciones psicométricas de ciertos inventarios de personalidad e intereses hace pensar que está próxima una mejora general en la evaluación de este tipo. (Aiken, 1996)

La probabilidad de que pruebas como el MMP1 o el Rorschach sean reemplazadas es baja, debido a su flexibilidad para adaptarse a situaciones cambiantes y a la presencia de decididos seguidores como Butcher o Exner. Se vislumbra el surgimiento de nuevas pruebas pero éstas forzarán una revisión continua de otras pruebas tradicionales. Autores como Weiner (1983) y Matarazzo (1992) han vaticinado un futuro brillante para las pruebas sobre la base de una investigación continua de los instrumentos (Aiken, 1996) Un campo promisorio es la investigación de respuestas fisiológicas asociadas a la evaluación de la inteligencia que más que sustituir a las pruebas de lápiz y papel, pueden complementarlas haciendo más precisa la medición. (Aiken, 1996)

Las pruebas y las evaluaciones psicológicas seguirán creciendo y formarán parte importante de la Psicología aplicada durante el próximo siglo. Cualquiera que sea el futuro de la evaluación psicológica, existirá la necesidad continua de evaluar la efectividad de los instrumentos y procedimientos psicométricos. Entonces, como ahora, la cuestión de la validez será la parte medular: los métodos de evaluación psicológica ¿cumplen sus propósitos?. En la actualidad, las pruebas ya no son la forma de subsistir de la Psicología pero siguen y seguirán siendo parte importante de esta. (Aiken, 1996) Los tests de personalidad han sido usados muy ampliamente por más de 50 años y hay suficiente evidencia que indica que hay un sitio para la evaluación psicológica, especialmente de la personalidad, en los lugares de trabajo. (Furham, 1992)

Además de las preocupaciones éticas y morales acerca de las pruebas de personalidad, las preguntas técnicas sobre lo que miden esas pruebas y si vale la pena medirlo así como la forma de interpretar y aplicar mejor los resultados, han recibido mucha atención legal y profesional durante las últimas décadas. No sólo existe una necesidad de pruebas de personalidad más confiables y válidas sino de mejoras en las bases teóricas de estos instrumentos y de los criterios con base en los que adquieren validez. Ninguno de los métodos disponibles para evaluar la personalidad son por completo satisfactorios para tomar decisiones en el campo laboral aunque una combinación de ellos aumenta la efectividad. Está claro que la solución para el problema de la validez radica en mejorar la investigación y el desarrollo, pero esos esfuerzos deben realizarse con una actitud socialmente responsable y con respeto de los derechos de los individuos.

Los usuarios de las pruebas deben poseer una mejor comprensión de la estadística y otros aspectos del diseño y propiedades psicométricas del test. Incluso en los casos en que la prueba resulta satisfactoria para los propósitos establecidos, es importante llevar registros de las pruebas, de indicadores de éxito o fracaso en la predicción, entre otros, pues a largo plazo esta información es en extremo útil para verificar la validez de la prueba. (Aiken 1996)

En cuanto al uso específicamente del Cuestionario 16 PF (versión de 1970 que es la que circula en México), en mi opinión es una prueba que no se conoce en toda su amplitud por lo que con frecuencia se desaprovecha información muy valiosa. Tiene muchas ventajas psicométricas pero la falta de adaptación a México es su principal desventaja. Se ha mencionado que la normas no han sido revisadas y no existen a la fecha baremos para poblaciones específicas como solicitantes de empleo o grupos ocupacionales; asimismo nunca se ha contado con normas para las formas B, C y D.

En esta prueba, el atraso fue tal que ya nos alcanza la nueva versión y sería más conveniente el adaptar esta a México, replicando los estudios psicométricos y obteniendo las normas para nuestro país. En España ya se ha adaptado la nueva versión del 16 PF-5 pero la traducción de una prueba a nuestro idioma no significa adaptación a nuestro país.

El MMPI-2 ya se publicó en México. Aunque se han hecho algunos estudios de adaptación, por la misma naturaleza del instrumento y su alcance, se requiere de investigación adicional que verifique sus propiedades en nuestro país.

Una alternativa viable es el Cuestionario de Personalidad Ocupacional, OPQ, que es una prueba de reciente creación, diseñada para el campo laboral, con estudios de adaptación a México aunque también requiere de otros estudios que constaten su utilidad en nuestro país.

3. CONCLUSIONES

El propósito de este trabajo fue el describir un método para derivar una escala de validez de un inventario factorial ampliamente usado en evaluación de personal en México, el Cuestionario 16 PF de Catell. Para contar con un marco de referencia, se hizo una somera revisión de las teorías de la personalidad y se encontraron hallazgos muy interesantes: la revisión hecha por Pedersen y cols. (1992 en Colom, 1995) y el estudio de Coan (1968 en Fernandez-Ballesteros, 1992), ambos importantes por las conclusiones a las que llegaron pero sobre todo, por el proceso utilizado para analizar las teorías. Coan utilizó análisis factorial para identificar dimensiones comunes a 54 teorías psicológicas. Pederson et al estudiaron 13 teorías para elegir la más apropiada para la evaluación de diferencias individuales en escenarios laborales según criterios valiosos por sí mismos.

Se refirieron algunos modelos de evaluación y se expusieron dos enfoques importantes en este campo: el proyectivo y el psicométrico. Se revisaron las teorías de Guilford y de Eysenck, quienes con Catell forman la triada de los principales factorialistas cuyos hallazgos permiten "iluminar o al menos reducir la oscuridad de la caja negra", en un intento por comprender básicamente a la Persona desde el enfoque de la Psicología Diferencial.(Colom 1995)

Asimismo, se relacionaron diferentes técnicas de evaluación con aspectos de costo, tiempo, especificidad y especialización del psicólogo. Se describieron los inventarios de personalidad como instrumentos psicométricos que forman una importante categoría. Se expusieron las fuentes de error en la evaluación de personal, se revisaron los conceptos acerca del falseamiento de las respuestas frente a los inventarios de personalidad y las escalas de validez, en general, como una alternativa para enfrentar esta importante limitación, para terminar con la descripción del método propuesto por Winder et. al. para construir una escala de distorsión motivacional útil sobre todo al usar el Cuestionario 16 PF de Catell en escenarios laborales.

En este trabajo, se presenta el enfoque psicométrico como una alternativa valiosa para la evaluación de personal; se han expuesto sus ventajas y limitaciones pero no ha sido la intención el plantear la disyuntiva de elegir entre instrumentos psicométricos o proyectivos pues no tienen que ser mutuamente excluyentes. Un psicólogo preparado en ambas técnicas puede exitosamente combinarlas cuando cuente con el tiempo para hacerlo y cuando el caso particular justifique el uso de técnicas proyectivas.

Los inventarios son la alternativa psicométrica para obtener información del personal pese a su susceptibilidad al falseamiento ya que, según Hough et al (1990 en Furham, 1992) este falseamiento no es importante y si éste es el costo de usar inventarios, el beneficio es a todas luces mayor, sin olvidar el hecho de que sus escalas de validez permiten detectar el tipo y la magnitud del falseamiento

La distorsión motivacional del 16 PF como un constructo ha sido identificada estadísticamente y validada por análisis factorial lo que a su vez comprueba la estructura de la personalidad planteada por Catell pues se encontró que las personas que simulan modifican sus puntajes en el sentido esperado por la teoría, por ej., se reduce el puntaje en ansiedad y se aumenta en fuerza del yo. Los resultados plantean a la Distorsión Motivacional DM como un constructo presente en sujetos bien adaptados, evaluados en una situación que propicia una modificación de sus respuestas para obtener algo. El procedimiento de Winder et al para derivar una escala de Distorsión Motivacional del 16 PF ha probado su utilidad al ser aplicado en diferentes países.

Pero la Distorsión Motivacional no es la principal fuente de falseamiento en la evaluación mediante tests psicométricos. El aspecto semántico y el nivel de comprensión verbal de los evaluados son aspectos que no se han investigado lo suficiente.

Una baja comprensión de la lectura puede dar puntajes altos en las escalas de validez pero en estos casos, estos puntajes son la punta del iceberg. La falta de comprensión o una comprensión errónea del significado de los reactivos afecta directamente la validez del instrumento. Cabe recordar que el nivel de comprensión de lectura no siempre se correlaciona con el grado de escolaridad.

3.1 Limitaciones de este trabajo y sugerencias para investigaciones relacionadas.

Este trabajo se circunscribe a la evaluación del falseamiento deliberado de los inventarios de personalidad. Se ha mencionado su utilidad en evaluación laboral, y por extensión y con base en la información revisada, se han hecho algunas propuestas de orden general o particular. Este tema está estrechamente vinculado con el proceso de selección de personal. A lo largo de este trabajo constantemente se hicieron referencias al mismo pero el estudio de este proceso no era el objetivo.

La selección de personal tiene que ver con el acoplamiento entre el puesto y la persona por lo que se necesita medir tanto al puesto como a la persona. Obviamente, la evaluación de la personalidad es una parte integral de la medición de la persona. (Kline, 1993b) pero la otra parte indispensable es la medición del puesto. En este trabajo no se tocó en absoluto el tema de la medición del puesto ni de la medición del desempeño, del cual derivarían los criterios para estimar la validez predictiva de las pruebas.

Asimismo, casi no se tocó el tema de las pruebas de capacidad en todas sus variantes, cuestión muy importante en la evaluación del personal.

Sugerencias para estudios posteriores

a) Directamente relacionados con el tema.

-Cualquier estudio empírico de aplicación o generación de escalas de validez en nuestro país

-Se requieren estudios lingüísticos de tipo semántico de los inventarios, sobre todo de aquéllos originalmente contruidos en otro idioma.

-Otra sugerencia es la investigación de la respuesta a los inventarios combinados con registros fisiológicos o de latencia de respuesta ante cada reactivo.

-Estudios de tipo correlacional entre pruebas de inteligencia y de comprensión verbal por un lado, con escalas de validez de inventarios, por el otro.

b) Indirectamente relacionados con el tema.

-Son urgentes los estudios de adaptación de pruebas extranjeras a nuestro medio. Así, el campo está abierto a normalizaciones, confiabilizaciones, validaciones.

-En las organizaciones laborales, el proceso de evaluación persigue propósitos específicos: una decisión de contratación o rechazo, de capacitación, de promoción, etc. pero rara vez los resultados de la evaluación se proporcionan a los evaluados con fines de autoconocimiento. Es necesario desarrollar modelos de retroalimentación de los resultados de una evaluación laboral que conduzcan a intervenciones no laborales, de prevención o de tratamiento en su caso.

-Un campo promisorio es el estudio de la conducta no-verbal aplicada a la evaluación de personal.

-Un tipo de pruebas que merecería atención especial por las circunstancias que se están viviendo en las grandes y medianas ciudades de nuestro país son las encaminadas a evaluar la honestidad: los tests de integridad.

-Y he dejado para el último el aspecto de mayor importancia: los estudios encaminados a la construcción de instrumentos propios que desde su diseño estén enfocados a las características de nuestra población.

* * *

El proceso irreversible y acelerado de globalización que estamos viviendo en todas las áreas de la actividad humana obliga a mantener vigente las técnicas de evaluación psicológica. Si hubo épocas en que este proceso era una experiencia local y temporal, actualmente ya no lo es.

No se puede uno sustraer al progreso de la tecnología pero éste de ninguna manera debe darse sobre la base de un detrimento de los avances científicos. Las tan trilladas cuestiones de validez, confiabilidad, normalización, etc. no pierden vigencia sino que ganan importancia ante la inminente expansión de los servicios de evaluación gracias a la computación. Es tarea y responsabilidad personal y social del psicólogo especialista el prepararse sistemática y concienzudamente para asesorar a la organización para la cual trabaja en la elección de los métodos de evaluación, incluyendo los programas computarizados, que mejor cubran las necesidades específicas de tal organización a corto, mediano y largo plazo así como el dignificar la actividad de evaluación psicológica, trabajando con ética, respeto y profesionalismo.

Un proyecto es un evento que tiene un principio y un fin. El camino para llegar al fin fue gratificante en sí mismo.

BIBLIOGRAFIA.

- Aiken, Lewis (1996) Tests Psicológicos y evaluación. México, Ed. Prentice Hall
- Arias Galicia Fernando (1996) Nuevos sistemas para la identificación de la personalidad ocupacional. SHL de México. Manual inédito
- Albert, Samuel et al (1980) Faking Psychosis on the Rorschach: Can expert judges detect malingering? *Journal of Personality Assessment*, 1980, 44-2, pp 115-119.
- Anastasi, Anne (1971) Tests Psicológicos. Madrid, Editorial Aguilar
- Beliak, Leopold (1990) TAT, CAT, SAT. México, El Manual Moderno
- Brown, F. (1980) Principios de Medición en Psicología y Educación. México, El Manual Moderno
- Buss Allen y Poley, Wayne (1979) Diferencias individuales, rasgos y factores. México, El Manual Moderno
- Catálogo de pruebas psicológicas. España, TEA Ediciones, 1995
- Catálogo 1997 de pruebas psicológicas. México, Ed El Manual Moderno
- Catell, Raymond, Eber y Tatsuoka (1988) Handbook for the sixteen personality factor questionnaire 16 PF. U.S.A. IPAT
- Catell, Raymond, Eber y Tatsuoka (1980) Cuestionario de 16 factores de la personalidad. Manual, instructivo y tablas. México, El Manual Moderno
- Catell, Raymond y Kline, Paul (1982) El análisis científico de la Personalidad y la Motivación. España, Ed. Pirámide
- Catell, Heather B. (1993) "Lo profundo de la personalidad" México, Ed El Manual Moderno
- Christiansen et al. (1994) "Correcting the 16 PF for faking: Effects on criterion related validity and individual hiring decisions. *Personnel Psychology*, Vol 47 (4) 847-860
- Colom Marañón, Roberto (1995) Inteligencia y Personalidad. Madrid, Editorial Pirámide
- Cueli y Reidl (1972, 1990) Teorías de la personalidad. México, Ed. Trillas.

- Editorial Rey. (1990) Diccionario ilustrado Vox latino-español. México.
- Fernández-Ballesteros, Rocio. (1992) Introducción a la Evaluación psicológica. Tomos I y II. España, Ed. Piramide
- Furham, Adrian. (1992) Personality at work. The role of individual differences in the workplace. London and New York, Routledge
- Groth-Marnat Gary (1990) Handbook of Psychological Assessment. New York, Chichester, etc. John Wiley & sons
- Hall, Calvin y Lindzey Gardner (1991) La teoría de la personalidad. México, Paidós
- Hall, Calvin y Lindzey Gardner (1984) La teoría factorial de la personalidad. México, Paidós
- Irvine, Michael J and Gendreau, Paul (1974) Detection of the fake "good" and "bad" response on the Sixteen personality factor inventory in prisoners and college students. *Journal of Consulting and Clinical Psychology*. Vol. 42, no. 3, p. 465-466
- Kirsón, Samuel y O'Dell, Jemy (1989) 16 PF, Guía para su uso clínico. Madrid, Publicaciones Teú
- Kline, Paul (1985) La Personalidad. Teoría y Medida. Madrid, Ed. Fundamentos
- Kline, Paul (1993) The Handbook of Psychological Testing. London, N.Y., Routledge
- Kline, Paul (1993b) Personality. The Psychometric View. London, Routledge
- Krug, Samuel E (1978) Further evidence on 16 PF Distortion Scales. *Journal of personality assessment*, 42, 5, pp 513-518
- Lanyon y Goodstein (1977) Evaluación de la Personalidad. México, El Manual Moderno
- Magnusson, David (1975) Teoría de los tests. México, Ed. Trillas
- Megargee, Edwin (1980) Métrica de la Personalidad. Vol. 1. México, Ed. Trillas
- Mischel, Walter (1988) Teorías de la personalidad. México, McGraw Hill
- Moms, Charles G. (1992) Psicología, un nuevo enfoque. México, Prentice Hall

- Nunnally, Jum C. (1982) Teoría Psicométrica. México, Ed. Trillas
- Núñez, Rafael (1994) Aplicación del MMPI a la psicopatología. México-Santafé de Bogotá. El Manual Moderno
- Rabin Ed (1968) projective Techniques in personality Assessment: a modern introduction. New york. Spenger Publishing.
- Rivera, Ofelia (1991) Interpretación del MMPI en Psicología Clínica, Laboral y Educativa. México, Ed. El Manual Moderno
- Sackett, Paul R. & Wanek, James E. (1996) New developments in the use of measures of honesty, integrity, conscientiousness, dependability, trustworthiness, and reliability for personnel selection. *Personnel Psychology*, 49, pp787-829
- Seisdedos, N. (1985) 16 PF. Monografía Técnica. España, TEA ediciones.
- Seisdedos, N. (1995) 16 PF-5. Manual. España, TEA ediciones
- Reader's Digest Ed (1979) Gran Diccionario enciclopédico ilustrado. México
- Thomton, George C. and Gierash, Paul F (1980) Fakability of an empncial derived selection instrument. *Journal of Personality Assessment*, 44, 1, pp 48-51
- Unbe Prado, Felipe (1996) "Factores sociodemográficos y su relación con habilidades de comunicación organizacional en alumnos con el manejo de un Assessment Center". Tesis de maestría inédita. Facultad de Comercio, UNAM.
- Valencia García, Ma del Rocío. (1996) Detección de perfiles de sujetos simuladores y sujetos honestos a través de las escalas del MMPI-2. Tesis de licenciatura inédita. Facultad de Psicología. UNAM
- Vernon, Philip E. (1982) Inteligencia, herencia y ambiente. México, Ed. El manual moderno