

LOS TESTS*

Mariano Yela

1. Definición y descripción general

La palabra inglesa *test* significa, como nombre, reactivo y prueba, y como verbo, ensayar, probar o comprobar. Se deriva del latín *testis*, cuya raíz figura en palabras españolas como *testigo*, *testimonio*, *atestiguar*, etc. Ha sido internacionalmente adoptada para denotar un cierto tipo de examen muy usado en Psicología y Pedagogía.

Un test es, en efecto, un *reactivo* que aplicado a un sujeto *revela y da testimonio* de la índole o grado de su instrucción, aptitud o manera de ser.

Así por ejemplo, un test de inteligencia consiste en un conjunto de preguntas, problemas o tareas cuya solución depende de las capacidades intelectuales del sujeto. Cada uno de los problemas preguntas o tareas se llama *elemento* o *item del test*. La primera condición que debe cumplir todo test es, naturalmente, que las respuestas suscitadas por sus elementos revelen la característica que desea apreciar o medir. Una vez demostrado ésto, se aplica el test a un grupo representativo de la población de sujetos en que se va a usar el test. A este grupo se le da el nombre de *grupo normativo*. Los resultados obtenidos por el grupo normativo, estadísti-

camente analizados, constituyen las *normas*, con las cuales hemos de comparar, en lo sucesivo, la puntuación que en el test alcance cualquier individuo de esa población para juzgar así del grado de su inteligencia o del rasgo que mida el test. Este proceso mediante el cual se fijan las normas de un test se llama *tipificación* del test.

Los tests de este tipo se llaman *psicométricos*, porque su objeto es medir o determinar el grado de algún rasgo mental.

Los test psicométricos son muy numerosos y variados. Para dar una idea de ellos presentamos a continuación algunos de los elementos o tareas que más se usan en su construcción.

Los elementos más corrientes en los tests de inteligencia son los siguientes:

Opuestos: El sujeto tiene que elegir los términos opuestos a una lista de palabras. Por ejemplo: *Alto...* grueso, pequeño, bajo, corto. En este ejemplo se pide que el sujeto subraye la palabra opuesta a alto. Este tipo de elemento, como todos los que se dan a continuación, se presta a valoración objetiva, fácil y rápida. Son susceptibles de ordenarse por su grado de dificultad y presentarse en varias formas y con diversas modificaciones. De tipo parecido son los elementos formados

* Yela, M. (1987). *Introducción a la teoría de los tests*, (pp. 3-29). Madrid: Facultad de Psicología, Universidad Complutense de Madrid.

por *sinónimos* y términos que guardan entre sí diversas clases de *relaciones*.

Analogías: Por ejemplo: *Rojo* es a sangre como amarillo es a ... papel, limón, bandera, ladrillo. Se pueden presentar analogías de este tipo con palabras, dibujos, personas, etc.

Palabra que sobra: Por ejemplo, subráyese la palabra que sobra entre las cinco siguientes: gris, marrón, amarillo, útil, rojo.

Serie de números: Por ejemplo, 1, 10, 2, 11, 3, 12, __, __. La serie tiene una ley. Hay que averiguarla y escribir sobre las rayas los números que siguen.

Vocabulario: Se presenta una serie, previamente experimentada y estudiada, de palabras y se pide al sujeto que diga lo que significan. Las respuestas se califican de distinta manera según su grado de abstracción.

Información: Consiste en una serie de preguntas sobre cultura general y datos supuestamente accesibles a todos los sujetos de la población en que se usa el test.

Frases incompletas: Es uno de los mejores elementos para apreciar la inteligencia. Consiste en un texto al que faltan algunas palabras que el sujeto tiene que averiguar y escribir en los espacios en blanco. Por ejemplo: El ___ brilla siempre, pero las nubes lo ___ algunas veces.

Otros elementos muy usados son: *Frases desordenadas*, que el sujeto tiene que comprender rápidamente; *proverbios*, cuyo sentido tienen que apreciar; *juicios prácticos*, o decisiones que es razonable tomar en situaciones que se plantean; *absurdos*, que el sujeto debe señalar en las escenas o sucesos que se le presentan; *memoria* de números, de palabras, de ideas, de formas, etc.; *interpretación* de párrafos, láminas, etc.

Todos estos elementos y otros similares reclaman respuestas de tipo *verbal* o *simbólico*.

Existen otros elementos de distinto tipo, que requieren respuestas gráficas y procesos *no verbales* como la realización de ciertos *dibujos*, la *comparación estética* de rostros, el manejo de *relaciones espaciales*, el *trazado de líneas*, el itinerario correcto de un *laberinto* y otros muchos de este tipo.

Las dos clases de elemento hasta aquí consideradas —verbales y no verbales— se denominan en general *elementos impresos o de papel y lápiz* porque bastan estos dos materiales para aplicarlos y resolverlos.

Hay otros elementos que exigen que el sujeto manipule trozos de madera, papel, metal, etc., o aparatos diversos, para resolver la tarea. Así hay tests que requieren formar un *rompecabezas*, reproducir un modelo con *cubos coloreados*, *repetir un ritmo* o un recorrido golpeando con un trozo de madera sobre cierto objeto o conjunto de objetos, *clasificar* objetos por su color, forma, volumen, peso, etc., o de todas las maneras que el sujeto sea capaz, *colocar trozos geométricos en tableros* con huecos especiales, etc. Todos ellos se llaman *elementos y tests manipulativos*. Se usan preferentemente en tests de las llamadas inteligencia práctica o concreta, habilidad manual o mecánica, en numerosos tests de aptitud profesional y con sujetos analfabetos, deficientes físicos o mentales, o con niños muy pequeños.

Los tests de personalidad utilizan elementos de otra clase, generalmente preguntas sobre *aspectos de la conducta* del sujeto, sobre sus *preferencias y aversiones*; sobre sus *actitudes y sentimientos*, sobre sus *tendencias e inclinaciones*, sobre sus *reacciones emotivas*, etc. Otros tests de personalidad utilizan el *trazado de líneas* y analizan el valor expresivo de los trazos, o bien se sirven de otros movimientos expresivos como el *gesto*, el *canto*, la *palabra*, el *paso*, la *escritura*, etc.

LOS TESTS

El método general al que se ajustan todos estos tests requiere:

a) Hallar elementos reveladores de la característica que se desea medir.

b) Plantear estos elementos de manera que puedan ser valorados cuantitativamente.

c) Averiguar en un grupo normativo las reglas para interpretar las diversas puntuaciones posibles.

Este es el tipo clásico de *tests psicométrico*.

Pero hay otros tests, aparecidos más recientemente y cada día más usados, que reciben el nombre de *tests proyectivos*. Pretenden asimismo revelar aspectos de la personalidad. Pero su objeto no es tanto la medida, como el diagnóstico. Sus elementos suelen ser *palabras, colores, manchas, material plástico, fotografías, films*, etc., que se presentan al sujeto para que, sujetándose a normas muy amplias, reacciones ante ellos en la forma y contenido que desee.

Un test proyectivo muy conocido consiste, por ejemplo en una serie de *escenas con figuras humanas*. El sujeto tiene que decir qué representa cada una y contar una historia acerca de la situación, los sentimientos de los protagonistas y el desenlace del asunto. Diferentes sujetos perciben en la misma escenas aspectos diferentes, con diferente significación, y los explican de diferente manera. Esas diferentes maneras de percibir y explicar reflejan las notas peculiares de cada sujeto. Su personalidad se *proyecta* en su *interpretación*. En estos tests es también preciso averiguar la significación de las respuestas en un grupo normativo si bien esto no será aquí fuente de reglas cuantitativas, sino de orientaciones para el diagnóstico.

En resumen, podemos concluir que *un test es una situación problemática, previamente dispuesta y estudiada, a que el sujeto ha de responder siguiendo ciertas instrucciones y de cuyas respuestas se estima, por*

comparación con las respuestas de un grupo normativo, de calidad, índole o grado de algún aspecto de su personalidad.

2. Origen y desarrollo del método

Los intentos de apreciar con algún rigor las diferencias psicológicas entre los individuos son muy antiguos. Varios modos de conseguir este propósito son indicados por los autores clásicos como Platón, Hipócrates, Teofrasto y Galeno y de un modo más concreto por el español Huarte de San Juan en el siglo XVI. Uno de los primeros psicólogos que elaboran y aplican procedimientos prácticos para este menester es el francés Seguin, hacia la mitad del siglo XIX. Seguin ideó varios tests manipulativos para la educación de los débiles mentales. Alguno de ellos se usan todavía como tests de inteligencia, por ejemplo, el *Tablero de Formas de Seguin*. Pero solo a fines del siglo pasado comienzan a usarse los tests de modo sistemático. Desde entonces su desarrollo ha sido progresivo y amplio; en él pueden distinguirse los cinco períodos que a continuación describimos brevemente.

A) Primer período: Tests sensoriales y motores

El primer período se extiende aproximadamente desde 1882 hasta 1905. En 1882 el inglés Galton, funda el *Laboratorio Antropométrico* de Londres. Su propósito principal es el estudio de la herencia de las aptitudes físicas y mentales. Con este fin incorpora a su laboratorio numerosos tests, la mayoría ideados por él mismo. Galton cree que la inteligencia puede ser convenientemente apreciada a través de la actividad sensorio-motora. Sus tests, y la mayoría de los que aparecen hasta finales del siglo, son, pues, de carácter *antropométrico, sensorial y motor*. La figura más representativa de este período

es quizá el americano J. Mc. K. Cattell, que introdujo por vez primera la palabra *test* en su artículo «Mental tests and measurements» aparecido en 1890 en la revista *Mind*. Cattell ideó y aplicó numerosos tests de este tipo en la Universidad de Columbia, de New York. En Alemania Munsterberg, Kraepelin y Ebbinghaus inician su aplicación. En Francia, lo hacen Binet, Henri y Toulouse.

Pero los resultados no fueron satisfactorios. En 1894 Gilbert halló una correlación casi nula entre varios tests de este tipo y el rendimiento escolar. Wissler, en 1901, confirma estos resultados. En Alemania Kraepelin y Oehrn, por su parte, y Ebbinghaus por otra, y en Francia Binet, empiezan a usar, en la última década del siglo, tests de procesos superiores, como asociación, frases incompletas, comprensión, etc. En 1889 Binet demuestra que los tests de procesos superiores distinguen mejor la capacidad escolar de los alumnos que los tests de tipo Galton.

Las características de este período son, pues:

- 1º. El predominio de tests sensoriales y motores.
- 2º. La comprobación de su ineficacia como medida de la inteligencia.
- 3º. El uso incipiente de tests de procesos más complejos.

B) Segundo período: Escalas de Inteligencia

Alfredo Binet, es, más que ningún otro, el creador de los tests de inteligencia. Durante los últimos años del siglo XIX, los intentos de discriminar el grado de inteligencia de los individuos por medio de tests, resultaron en general fallidos. Binet comprendió que el efecto no estaba en el método, sino en el material empleado. Se usaban tests de procesos elementales. Había que introducir elementos que estimulasen las funciones psicológicas complejas superio-

res. Este es uno de sus aciertos. El otro gran acierto fue relacionar los tests con la edad, formando la primera *escala de inteligencia*. Binet observó que, al menos durante los años escolares, los niños van teniendo más capacidad mental cuanto mayores son. Basado en esta sencilla observación ideó una serie de pruebas de fácil y rápida aplicación, las ensayó en las escuelas de París, y determinó para cada prueba el tanto por ciento de niños que la resolvían a cada edad. Así llegó a averiguar los tests que eran característicos de cada edad. Si un test es correctamente resuelto por el niño medio de 12 años, pero no por los niños de edades inferiores, ese test representa la capacidad mental típica de los doce años, o sea, la *Edad Mental* (E.M.) de 12 años. Ordenando los tests por edades sucesivas formó pues, una escala mental. La primera fue publicada en 1905, pero fue revisada y completada en 1908, y es esta segunda la que alcanzó enseguida fama mundial. Binet seleccionó de 3 a 8 tests para representar cada edad de los 3 a los 13 años. Cada test equivale a un cierto número de meses. La puntuación total del niño es el número de años y meses de edad mental que ha obtenido en sus respuestas. La escala está preparada de tal modo que el niño medio de cada edad, obtiene una edad mental (E.M.) igual a su edad cronológica (E.C.) Si un niño de 7 años obtiene una E.M. de 9 años, es un niño mentalmente adelantado, si su E.M. es de 5 años, está retrasado mentalmente.

La escala de Binet tuvo un éxito extraordinario. Se comentó, criticó, tradujo y adaptó paulatinamente a todos los idiomas principales. De las múltiples revisiones del test de Binet, la más famosa, por mejor construida y tipificada, es la del psicólogo americano Terman, llevada a cabo en la Universidad de Stanford en 1916, y revisada en 1937 y en ulteriores ocasiones. El test Stanford-Binet, de Terman, ha sido adopta-

LOS TESTS

do, también, en casi todos los países. En España lo han hecho el Dr. Germain y la Srta. Rodrigo.

Una modificación importante de este tipo de escalas fue introducida por Yerkes y Bridges, en 1915. Consiste en valorar no por meses, sino por puntos. La puntuación final no es una E.M. sino simplemente el número de puntos obtenidos. Ahora bien, como la puntuación media de los individuos de una cierta edad cronológica representa esa edad mental, las puntuaciones totales se pueden convertir fácilmente en edades mentales, por medio de una tabla de correspondencias entre edades mentales y puntuaciones.

Para eliminar algunos inconvenientes que presentaban las edades mentales —por ejemplo, un año de edad mental significa mucho más en los primeros años del desarrollo que en los más avanzados—, se propuso usar el cociente entre la E.M. y la E.C. Esta medida fué propuesta, independientemente al parecer, por el americano Kuhlmann y el alemán Stern, en 1912. Terman aceptó esta medida y generalizó el uso del *cociente intelectual* (C.I.) El C.I. es, pues el cociente entre la E.M. y la E.C. Para evitar decimales, se suele multiplicar por ciento. De este modo el C.I. indica el tanto por ciento que la E.M. de un individuo es de su E.C.

Una E.M. de 10 años no dice nada acerca de la agudez mental de la persona, sino simplemente que tienen el nivel mental de un niño corriente de 10 años. Si su E.C. es de 14 años, el chico será torpe; si es de 7 años, el chico será listo, como indican sus C.I. que serían 71 y 143 respectivamente.

A partir de la escala de Binet y con modificaciones diversas, se han elaborado tests de este tipo. Quizá la más importante de las modificaciones es la escala "Wechsler-Bellevue", una de las mejores hasta la fecha para apreciar la inteligencia

general de los adultos. Las características generales de todos estos tests, son:

1º. Predominio de elementos que hacen referencia a procesos cognoscitivos superiores.

2º. Variedad de los elementos, procurando estimular con ellos las más diversas funciones mentales.

3º. Normas valorativas fundadas en la edad cronológica.

4º. Puntuación global indicativa de la inteligencia general del sujeto.

5º. Aplicación individual del test.

C) Tercer período: Tests colectivos

En 1917 aparece en América un nuevo tipo de test: el *test colectivo*. Hasta ahora los tests tenían que ser aplicados individualmente. Al entrar los Estados Unidos en la primera guerra mundial, el país no tenía ejército. Para crearlo rápidamente era preciso clasificar a los reclutas según sus capacidades. Los tests podían servir para esto. Pero hacía falta que pudieran ser aplicados en poco tiempo a miles de individuos. Un comité de psicólogos presidido por Yerkes y formado, entre otros, por Terman, Whipple y Goddar, todos ellos eminentes en este campo, prepararon una serie de tests cuyos elementos, de tipo análogo a los ya citados, van impresos en una o varias hojas, las cuales contienen además las instrucciones generales. El examinador solo tiene que hacer una breve explicación al grupo y procurar que el test se realice según las instrucciones. Este tipo de test se llamó "Alfa". Para los analfabetos o extranjeros se creó el test "Beta", en el cual los problemas eran formulados sin hacer uso de palabras, sino mediante figuras, dibujos, etc., y las instrucciones eran dadas por medio de mímica o con demostraciones prácticas. Estos tests se aplicaron a cerca de dos millones de hombres, sirviendo principal-

mente para rechazar a los ineptos y clasificar, según sus aptitudes, a los útiles para el Servicio Militar.

De este primer ensayo nacen y se difunden por todo el mundo numerosos tests colectivos. Su sistema de valoración es sencillo. Consiste, en general, en el número total de aciertos. Para interpretar estas puntuaciones se recurre, bien a normas cronológicas como E.M. y C.I., bien a normas estadísticas, como centiles o puntuaciones típicas.

Los tests del Ejército americano tuvieron otra consecuencia importante: generalizar el uso de *tests de manipulación*. Estos son tan antiguos como el mismo método de los tests. En los primeros años del siglo, los utilizó sobre todo el italiano Sancte de Sanctis. Todas las escalas de inteligencia incluyen elementos manipulativos. Pero son los tests del Ejército, principalmente los de tipo Beta, los que dan el impulso decisivo para la creación de tests y escalas de inteligencia, exclusivamente, manipulativas. El primer ejemplo lo constituye la *Escala de tests manipulativos* de Pintner y Paterson, que aparece en 1917 y ha experimentado varias revisiones posteriores. Contiene tests que requieren construir o completar ciertas escenas o figuras manejando bloques de madera con dibujos o formas especiales. Los tests son en general rompecabezas de dificultad creciente, por ejemplo, construir el *perfil* de una cara, un *maniquí*, un *barco*, una *yegua con su potro*, y otras escenas, o colocar ciertos bloques de formas geométricas diversas en los huecos de un tablero. Las tareas se valoran en función del tiempo y los errores.

Según este sistema, con variaciones más o menos grandes, se han construido numerosas escalas de «performance» o manipulación.

Este período introduce, pues, en el método de los tests, las siguientes novedades:

1º. Aplicación colectiva.

2º. Predominio de elementos exactamente cuantificables y mayor exactitud en los aspectos estadísticos del método.

3º. Mayor facilidad, rapidez, economía, y objetividad en la aplicación e interpretación.

4º. Generalización del uso de tests manipulativos.

D) Cuarto período: Tests de personalidad

En 1917 Woodworth preparó en los Estados Unidos la llamada *Hoja de datos personales*, en la cual, mediante preguntas sobre hábitos, intereses, actitudes, temores, etc., se pretendía llegar a un diagnóstico sobre la personalidad y problemas del sujeto. En 1921 el *Test Psicodiagnóstico* del psiquiatra suizo Rorschach. Es el primer test que pretende dar una visión comprensiva y dinámica de la personalidad. Es uno de los más usados actualmente. Consiste en unas manchas de tinta que el sujeto debe interpretar, contestando a la pregunta ¿qué ve usted en esta lámina? ¿a qué se parece esto? Como el estímulo que se presenta es ambiguo, la respuesta con que reacciona el sujeto depende, según Rorschach, de su manera de percibir, fundada asimismo en la estructura de su personalidad.

Siguiendo el tipo de cuestionario, el modelo proyectivo u otras técnicas más o menos objetivas, se han creado innumerables tests que pretenden revelar la personalidad o algún aspecto afectivo, temperamental, social, etc., de la misma.

Su características principales son:

1º) Extender el método de los tests al campo de la personalidad.

2º) Inclinarsé más bien al análisis cualitativo que a la medición.

3º) Crear un nuevo tipo de test: el proyectivo.

E) Quinto período: El análisis factorial

Hasta ahora hemos visto como nuevas ideas iban ampliando el campo de los tests. En el primer período se utilizan pruebas sensoriales y motores. En el segundo, desde 1905, aparecen tests individuales de inteligencia general. En el tercero, desde 1917, se emplean también tests colectivos de inteligencia. En el cuarto período, que se inicia aproximadamente, por los mismos años, pero que solo alcanza verdadera difusión en la segunda guerra mundial, se extiende el método al campo de la personalidad y se inventa un nuevo tipo de tests, el proyectivo.

Pero a través de todos estos cambios y ampliaciones persiste, sin resolver, un problema fundamental: ¿Qué es lo que miden o diagnostican los tests? ¿Cómo estar seguro de que miden los rasgos que los tests dicen medir y de que en efecto los miden? No basta con que los tests sean útiles y prácticos. Es preciso demostrar de alguna manera que aprecian o miden las capacidades y rasgos unitarios de la personalidad.

La consideración de este problema dió origen a una serie de técnicas que se conocen con el nombre de *análisis factorial*. Su creador Spearman, cuyos primeros trabajos aparecen ya en 1904, y cuya teoría se publica en 1927, en el libro "*The Abilities of Man*". El análisis factorial es un método para averiguar experimentalmente lo que un test mide o aprecia. Los trabajos de Spearman han sido continuados principalmente en Inglaterra y los Estados Unidos. Hoy sobresalen los nombres de Burt, Thomson y Thurstone. Este último, basado en los hallazgos del análisis factorial, ha preparado una serie de tests de "*Capacidades Mentales Primarias*" con los que pretenden apreciar las capacidades cognitivas, factorialmente

descubiertas y definidas, de que depende la actividad inteligente del individuo.

El análisis factorial ha influido decisivamente en el desarrollo del método de los tests. Respecto a los tests de inteligencia su influjo ha sido doble. Por una parte ha ofrecido una seria fundamentación teórica al uso de los antiguos tests de inteligencia general, mostrando que en la mayor parte de las funciones cognitivas intervienen un factor general (factor *g*), común a todos ellos y caracterizado por la abstracción y comprensión de relaciones. Los tests de inteligencia general están justificados en la medida en que aprecian el factor "*g*". Para ello se dispone hoy de tests más adecuados que los tradicionales, como son las *Matrices de Raven* y el test de Dominó. Por otra parte, los resultados del análisis factorial subrayan la importancia en el campo de la inteligencia de otras aptitudes cognitivas distintas del factor "*g*" y que es preciso tener en cuenta para apreciar suficientemente la inteligencia de los sujetos. No basta, pues, con designar la inteligencia general mediante un índice global, es preciso trazar, además, el perfil mental del sujeto según la puntuación que alcanza en diversos factores cognoscitivos y motores, como inteligencia verbal, numérica, espacial, memoria, percepción, etc.

En el campo de la personalidad emotiva y social el análisis factorial ha señalado como dimensiones básicas las correspondientes a la *madurez emotiva* y a la *introversión-extraversión* de los sujetos. Asimismo se han desarrollado numerosos cuestionarios y tests de intereses generales y profesionales a partir de los resultados obtenidos por el análisis factorial.

El desarrollo que ha alcanzado el método de los tests lo manifiesta el hecho de que solo en los Estados Unidos se aplicaron durante 1946, unos sesenta millones de tests a unos veinte millones de personas. El núme-

ro de aplicaciones, desde entonces, no ha hecho más que aumentar y multiplicarse.

3. Clasificación

Los tests son muy numerosos y variados. Para hacerse una idea de los principales tipos, es preciso clasificarlos. La clasificación que presentamos a continuación, sin pretender ser única o completa, tiene por objeto ofrecer un cuadro a la vez comprensivo y simple de los múltiples tests existentes.

Un test es un proceso, ajustado a ciertas *normas generales*, al cual *una persona* es sometida por *otra* utilizando cierto *material*, para medir o diagnosticar cierta *característica* y con un determinado *fin*. Podemos, pues, clasificar los tests atendiendo a:

- a) Normas generales del método.
- b) El fin que se persigue.
- c) Quién los da.
- d) Quién los recibe.
- e) El material que se emplea.
- f) La característica que pretenden estimar o medir.

A.- Clasificación por el método

Según el método general que siguen los tests, se clasifican en *psicométricos* y *proyectivos*. Es la división más importante y significativa. Los psicométricos se ajustan a normas cuantitativas, sus elementos son valorados independientemente unos de otros con una puntuación numérica, su resultado final es una medida. Los proyectivos se ajustan a normas principalmente cualitativas, sus elementos son valorados cualitativamente en relación con el test total, su resultado final es un diagnóstico. Los tests psicométricos se refieren a características psicológicas, concebidas como unidades más o menos independiente. Los proyectivos se refieren en general, a la personalidad total. Estas diferencias no son, sin embargo, absolutas, sino

relativas y variables. Hay aspectos proyectivos en los tests psicométricos, principalmente en los individuales, y hay aspectos cuantitativos en los proyectivos. Pero la diferencia es clara. Los psicométricos son más rigurosos y exactos, mas a costa, muchas veces, de prescindir de los aspectos más profundos y reales de la personalidad. Los proyectivos son más comprensivos, pero a costa, muchas veces, de caer en subjetivismos más o menos arbitrarios. Hay que tener en cuenta, sobre todo, que para aplicar los tests proyectivos se requiere una preparación teórica y una experiencia clínica mucho mayor que para utilizar los psicométricos.

B.- Clasificación por el fin

El fin de los tests es doble: *diagnosticar* y *medir*. Según esto, hay tests *cualitativos* y tests *cuantitativos*, división que coincide, aunque no por completo, con la de tests proyectivos y psicométricos.

Por otra parte, los tests pueden utilizarse como *método práctico* y como *método de investigación*. El test es fundamentalmente un método práctico. Su objeto es señalar la presencia, calidad o grado de algún rasgo. Pero desde principios de siglo, y sobre todo, en los últimos años, se han utilizado, junto al análisis factorial, en la investigación de las funciones psicológicas. No hay pues una distinción absoluta entre tests y experimento psicológico. Pero si cabe afirmar que, en general, el test se refiere a las características de las personas, mientras el experimento se refiere a las características de los fenómenos o procesos.

C.- Clasificación por el examinador

Los tests se dividen, según este criterio, en *personales* e *impersonales*. Son personales aquellos en que la personalidad o conducta del examinador influye en los resultados.

LOS TESTS

En rigor, todos los tests son personales. Lo son muy especialmente los proyectivos, en los que el examinador ha de procurar liberar al sujeto de toda presión racional o emotiva que le impide expresarse libremente, y, por tanto, adaptar su manera de proceder a las condiciones únicas de cada sujeto. Lo son asimismo, aunque en menor grado, los psicométricos individuales, en los que el examinador influye con su manera de hacer preguntas, su gesto, su actitud general. En menor grado lo son los colectivos, donde el examinador se limita a establecer un "rapport" afectivo adecuado con el grupo. Estos últimos tests pueden considerarse, pues, como impersonales. Conviene tener muy en cuenta este aspecto de los tests, a saber, la influencia de la personalidad y conducta del examinador, para contar con ella y emplearla consciente y oportunamente cuando convenga, y tratar de limitarla o hacerla, en lo posible, uniforme, cuando el carácter del test así lo reclame.

D.- Clasificación por el sujeto

Se llama sujeto al que recibe el test. Atendiendo al sujeto, los tests se dividen en *individuales* y *colectivos*. Según se puedan aplicar tan solo a un sujeto o varios a la vez. Los tests colectivos son fáciles de aplicar o interpretar. Los puede aplicar toda persona que tenga una mínima preparación psicológica. Los individuales exigen una preparación muy especial en la ciencia psicológica y una larga experiencia práctica.

Se pueden hacer muchas otras clasificaciones de los tests, atendiendo al sujeto que los recibe. Las de mayor interés son las siguientes:

Tests destinados a sujetos *normales* y tests para *anormales* ya sean físicos o mentales.

Tests destinados a la *población general*, o tests para *grupos especiales* (de cierto

nivel cultural, social, escolar, profesional, racial, etc.).

Y, finalmente, tests destinados a sujetos de *todas las edades* y tests destinados a alguna *edad especial*. Conviene saber, sobre todo, que existen tests especiales para los *primeros meses y años de la vida*, y tests especialmente preparados para *adultos*.

E.- Clasificación por el material

Entre las muchas clasificaciones posibles, la de mayor valor práctico es, a nuestro parecer, la que distingue entre test *impresos* y tests *manipulativos* (véase sección 1).

F.- Clasificación por el rasgo que aprecian o miden

Según este criterio, los tests se pueden clasificar según el cuadro siguiente:

a) Tests de rendimiento:

- Escolares
- Profesionales

b) Tests de aptitud:

- Psicológica:
 - . Inteligencia general
 - . Factores cognoscitivos
 - . Aptitudes especiales
- Práctica:

- . Escolares
- . Profesionales

c) Tests de personalidad:

- Por el fin:
 - . Sintéticos: Personalidad total.
 - . Analíticos: Actitudes, Intereses, Valores, Emotividad, etc.
- Por el medio:
 - . Subjetivos
 - . Expresivos
 - . Proyectivos
 - . Objetivos
 - . Situacionales

a) Tests de rendimiento

Su objeto es medir con precisión y objetividad los conocimientos o el rendimiento actual de un sujeto en relación con los que, en una determinada materia, han mostrado poseer los sujetos de su población.

Así, hay tests sobre conocimientos en las diversas materias escolares y académicas; son como los exámenes tradicionales, pero sistematizados, objetivos y tipificados. Existen, asimismo tests para apreciar la competencia profesional de un individuo en cierta ocupación, siempre en relación, claro está, con un grupo normativo de la misma.

b) Tests de aptitud

Llamamos aquí *test de aptitud*, a todo aquél que se refiere a las capacidades cognitivas y psicomotoras potenciales del sujeto y no a su actual grado de conocimientos o competencia. Estos tests pueden definirse, bien en relación con la estructura psíquica del sujeto, y los llamaremos tests de *aptitud psicológica* como son los de inteligencia, de imaginación, de coordinación psicomotora, etc.; o bien en relación con las exigencias de una determinada situación externa, en cuyo caso los llamaremos tests de *aptitud práctica*, como son los de aptitud para las matemáticas, para la ingeniería, para tornero, o impresor, o aviador, etc.

Entre los tests de aptitudes psicológicas destacan en primerísimo lugar los llamados de *inteligencia general*, de los cuales ya hemos dado una idea más arriba, y en segundo lugar —aunque cada vez se les concede mayor importancia— los tests que miden las diferencias características de esa inteligencia general o, como hoy suele decirse, los tests de factores *cognoscitivos* y de *aptitudes primarias*, como son los de *razonamiento abstracto*, *comprensión verbal*, *fluidez verbal*, *facilidad ideativa*, *razonamiento cuanti-*

tativo, *imaginación espacial*, *facilidad y rapidez perceptiva*, *síntesis perceptiva*, *memoria lógica*, *memoria inmediata*, etc.

Existen, además, tests de capacidades especiales, como son, por ejemplo los de *destreza manual*, *rapidez motora*, *habilidad mecánica*, *apreciación estética*, etc., además de los tests tradicionales de *sensibilidad y precisión motora*.

Con respecto a las aptitudes prácticas existen numerosos tests y grupos de tests (batería) para apreciar la disposición de los individuos hacia los distintos trabajos y profesiones. En la enseñanza se usan tests de *aptitud para las matemáticas*, para la *física*, para los *idiomas*, etc., y tests más generales que distinguen entre los sujetos más aptos para el estudio de disciplinas académicas y aquellos que tienen más aptitud para tareas concretas, activas, mecánicas o sociales. En la Industria, el Ejército y otras empresas oficiales y privadas, se utilizan baterías de tests para seleccionar al personal más apto. Así, hay baterías de tests para seleccionar distintos tipos de aviadores, oficinistas, mecánicos, etc.

c) Tests de personalidad

La personalidad es la concreta realización de la persona: su manera de ser y comportarse. Abarca, pues, todo el psiquismo con sus perspectivas pretéritas y futuras según se articulan en un presente dinámico. En su aspecto psicológico se forma y modifica a lo largo de la vida. No es ésta sin embargo la significación que suele tener la palabra *personalidad* cuando se usa en relación con los tests. En este contexto adopta por lo general un sentido más limitado. Suelen, en efecto, llamarse tests de personalidad aquellos que se refieren a aspectos no cognoscitivos de la conducta tales como los afectos, inclinaciones y actitudes. La personalidad, en este sentido, se refiere al tempe-

ramento carácter y hábitos sociales del sujeto. Los tests de personalidad pueden clasificarse según el fin que se proponen y según el medio que para ello utilizan. Por el fin, cabe distinguir los tests sintéticos de los analíticos. Los tests sintéticos pretenden apreciar la *estructura general de la personalidad*, sus *tendencias y problemas* más salientes. Suelen ser de carácter proyectivo como el Rorschach, el T.A.T., el Szondi, etc. Otros son de tipo analítico y aprecian aspectos parciales de la personalidad, entre estos últimos, existen tests sobre *intereses y preferencias* ya generales, ya con respecto a un tema especial, por ejemplo, el profesional; tests de *actitudes*, que versan, asimismo, o bien sobre actitudes generales del sujeto frente a la vida y se llaman tests de *valores*, o bien sobre su actitud con respecto a un tema más o menos limitado, por ejemplo, sistemas políticos, organizaciones sociales, profesiones, carreras, razas, religiones, etc.; tests de *rasgos especiales de carácter*, como sinceridad, puntualidad, cooperación, perseverancia, etc., y finalmente, tests de *rasgos temperamentales y emotivos*, como introversión, sociabilidad, equilibrio emotivo, agresividad, etc.

Los medios que se utilizan en los tests de personalidad son muy diversos. Ya hemos mencionado algunos. La mayor parte pueden clasificarse en: 1º) Subjetivos: como son los cuestionarios, en los que el sujeto ha de responder a preguntas más o menos explícitamente referidas a sus preocupaciones, hábitos, opiniones, sentimientos, etc., tales, los cuestionarios de Guilford y de Eysenck sobre la estabilidad emotiva y la introversión de los sujetos. 2º) Expresivos: en los que se estima la personalidad del sujeto a través del estudio de los movimientos expresivos del mismo, como acontece en el test meocinético de Mira y en los diversos intentos grafológicos. 3º) Proyectivos: de los que ya se ha hablado antes. 4º) Objetivos: en los que se

observa y mide algún aspecto de la conducta que previamente ha mostrado tener relación con determinadas características de la personalidad; así, el estudio de la respuesta psicogalvánica como índice de emotividad. 5º) Situacionales: en los cuales se observa la conducta del sujeto en una situación real, por ejemplo, el juego, la representación de un papel en un drama (psicodrama), su participación en una tarea de discusión de un grupo, etc.

4. El test como instrumento de medida

Todo test psicométrico debe reunir como mínimo, tres condiciones fundamentales: ser *válido*, ser *preciso* o *fiable* y estar *tipificado*. Para conseguirlo la elaboración de un test suele seguir el esquema siguiente:

- A) Definición del rasgo psicológico.
- B) Selección de elementos apropiados.
- C) Comprobación de la validez y precisión del test.
- D) Tipificación del mismo.

A.- Definición del rasgo

El creador de un test está siempre dirigido por una hipótesis de trabajo: su concepción de la característica que quiere medir. Esta hipótesis debe formularse con la mayor claridad, y precisión. Como, en general, se tiene un conocimiento muy precario y limitado de la estructura del psiquismo, la característica cuyo grado se desea revelar no suele concebirse como una capacidad o rasgo bien determinado, sino, simplemente, como un aspecto del psiquismo o de la conducta cuya medida, aunque no descubra la naturaleza intrínseca de este aspecto, permita conocer de antemano, con cierta probabilidad, cuál será la conducta del individuo en

otras ocasiones, al menos en cuanto se refiere a ese aspecto de su comportamiento.

El ideal de un test es, sin embargo, apreciar un rasgo unitario de la personalidad. Una de las mejores maneras de definir experimentalmente un rasgo unitario, para luego elaborar un test que lo aprecie o mida es, por ahora, el análisis factorial.

B.- Selección de los elementos

Una vez claramente definido el rasgo psicológico, se procede a inventar o reunir elementos que lo pongan de manifiesto. La clase de elementos que se elijan depende, en primer lugar, de la concepción que se tenga sobre el rasgo en cuestión. Es claro, sin embargo, que cualquiera que sea nuestra manera de concebir una característica psicológica, ésta puede manifestarse en un número infinitamente variado de situaciones o elementos. El problema consiste en elegir, para formar el test, los más útiles y significativos. En esta selección se utilizan algunas técnicas estadísticas auxiliares, sobre todo para comprobar la *validez* y la *dificultad* de los elementos.

Para asegurarse de que los elementos son válidos, se prepara un número considerable de ellos y se eligen, después de ensayarlos.

a) Los que tienen una correlación alta con el total, con lo que se asegura la unidad interna del test, es decir, que todos los elementos contribuyan a medir lo mismo.

b) Los que tienen alta correlación con un criterio externo de validez, o con una dimensión factorial previamente determinada. De este modo solo se retendrán elementos que contribuyan eficazmente al propósito del test, a saber, producir un cierto criterio o apreciar un rasgo estable del psiquismo.

Para averiguar la dificultad de los elementos basta ensayarlos en un grupo

representativo y hallar el tanto por ciento de individuos que pasan cada elemento. Cuanto mayor sea el tanto por ciento de individuos que pasan un elemento, tanto menor será su dificultad. Según estos tantos por ciento, u otros índices de ellos derivados, se pueden ordenar los elementos por su dificultad de la manera que mejor convenga.

C. Comprobación del test

Una vez reunidos los elementos más convenientes, según su validez y dificultad, se tiene construido el test. Antes de usarlo, sin embargo, es *absolutamente imprescindible demostrar*:

- a) Que es suficientemente válido, y
- b) Que es suficientemente fiable o preciso.

a) La validez del test

Un test que se utilice para seleccionar conductores es válido si distingue entre buenos y malos conductores. Un test de inteligencia es válido si distingue entre los torpes y los listos. Para comprobar la validez de un test de selección de conductores se halla, pues, su correlación con la eficacia profesional de éstos. Para comprobar la validez de un test de inteligencia se halla su correlación con el grado de inteligencia manifestado por los sujetos en su vida escolar, familiar, profesional, etc., En ambos casos es preciso, ante todo, fijar un *criterio externo* —esto es independientemente del test— con arreglo al cual se puedan clasificar los sujetos en la característica que el test pretende medir. Este criterio podría ser la destreza en el volante según es estimada por un grupo de jueces competentes, en el caso de los conductores. O los juicios de maestros, jefes, padres, amigos, etc., sobre la inteligencia de los sujetos, o quizá su éxito

LOS TESTS

escolar, o las estimaciones de clínicos que conocen bien al paciente, etc., en el caso del test de inteligencia. El test es válido si tiene una alta correlación con el criterio.

Se comprende que este método de validación no es perfecto, pues no lo son en general los criterios. Uno de los puntos más importantes en la elaboración de un test, es, pues, el estudio minucioso de los criterios de validación, para eliminar de ellos todo posible error.

De todos modos, esta validación es exclusivamente pragmática. Sólo indica que el test sirve para lo que se le destina. Pero no si revela algún rasgo significativo del psiquismo. Por consiguiente, para que los tests puedan perfeccionarse, es preciso, además de criterios externos que garanticen su valor práctico, *criterios psicológicos* que puedan aclarar el significado interno de los rasgos y funciones que los tests miden. Hacia este objetivo se orienta el análisis factorial. Así otro medio de estimar la validez de un test es hallar su correlación con una unidad funcional que, de antemano, se haya determinado factorialmente.

b) La precisión o fiabilidad de un test

Se llama precisión de un test al grado de constancia de sus medidas. Si al aplicar un test varias veces a un individuo, éste obtiene cada vez puntuaciones muy distintas, el test es impreciso, no podemos fiarnos de sus resultados. Si las puntuaciones son iguales o difieren poco, el test proporciona medidas precisas, fiables, dignas de confianza. La precisión suele estimarse mediante la correlación del test consigo mismo, o por algún otro procedimiento equivalente. Antes de utilizar un test como instrumento de medida es, pues, imprescindible que se averigüe su precisión para saber el grado de confianza que merece sus resultados cuantitativos.

c) Circunstancias que favorecen la validez y precisión de un test

Las principales son las siguientes: La definición cuidadosa del rasgo: su pureza factorial y la determinación rigurosa de criterios de validación. El uso de elementos claros, inequívocos, objetivos, ordenados, independientes, interesantes para el sujeto, que cubran bien el campo de dificultad del rasgo, aunque con predominio de elementos de dificultad media. La redacción de instrucciones breves y claras, adecuadas al nivel mental mínimo de los sujetos. La introducción de ejercicios preparatorios. La estricta observancia de las instrucciones por parte del examinador, y el establecimiento de un "rapport" adecuado entre el examinador y el sujeto.

En igualdad de circunstancias, se prefieren tests breves, baratos y de material sencillo, fáciles de aplicar, objetivos y rápidos en su valoración.

D.- Tipificación

Elaborado un test y comprobadas su validez y precisión, queda por determinar la significación de las puntuaciones. El resultado de un test es un número o *puntuación directa*. Con ella no sabemos si el sujeto es normal o no, ni la dirección y magnitud de su desviación en el caso de que no lo sea. Para averiguar la significación de una puntuación dada, es preciso compararla con las obtenidas por la *población* a que el sujeto pertenece. Averiguar y ordenar las puntuaciones de la población es *tipificar* un test. Como, en general, no es posible someter al test a *toda* la población, es preciso, ante todo, seleccionar una *muestra representativa* de la misma, a la cual llamamos *grupo normativo* porque de él se obtienen las *normas*. Una vez elegida la muestra, se aplica el test a todos sus miembros. Las

puntuaciones obtenidas se ordenan en forma de *escala cuantitativa*, que sirva para interpretar, en lo sucesivo, el resultado que obtenga un determinado sujeto.

Para ello es preciso que la escala tenga un punto de partida de significación conocida y fija, y que su unidad se mantenga constante a lo largo de la escala, o, por lo menos, varíe en forma conocida. Hay varias maneras de hacer esta ordenación y, por consiguiente, varios tipos de normas. Los principales son tres: normas cronológicas, normas centiles y normas típicas.

a) Normas cronológicas

Se han empleado casi exclusivamente con los tests de inteligencia. El procedimiento, en esencia, consiste en averiguar cuál es la puntuación media que cada edad obtiene en el test. La puntuación media de cada edad designa la *edad mental* correspondiente. Así, si la puntuación media de los niños de 8 años es, en un test de inteligencia, 26 puntos, en ese test 26 puntos significa 8 años de edad mental. Esta escala tiene dos graves inconvenientes. El primero es que su unidad no es constante. Un año de edad mental no es lo mismo a lo largo del desarrollo. Al parecer, y prescindiendo de diferencias cualitativas (que, no obstante, siempre son del máximo interés), la distancia mental de un año al siguiente va disminuyendo con la edad. Esto, si no se tiene en cuenta, puede inducir a errores. El segundo inconveniente es que en todos los tests de inteligencia actuales se llega a una *edad tope* a partir de la cual la puntuación media no aumenta con la edad. Por lo tanto, a partir de esta edad tope —de los 14 a los 18 años—, no tiene aplicación al concepto de edad mental.

Para eliminar el primer inconveniente se usa el *cociente intelectual*, o sea, la razón entre la edad mental y la edad cronológica, generalmente multiplicada por ciento para

evitar decimales. Para que una escala de cocientes intelectuales sea correcta, es preciso que cualquier C.I. dado signifique cuantitativamente lo mismo en todas las edades. Para ello es necesario que el C.I. medio de todas ellas sea siempre igual a 100, y que las distribuciones de C.I. de todas las edades, tengan la misma forma y la misma dispersión. Si se cumplen estas condiciones, al menos aproximadamente, la escala de C.I. indica con precisión la agudeza mental de los sujetos. Queda en pie, sin embargo, el segundo inconveniente. El C.I. no es aplicable, en un test de inteligencia, por encima de su edad tope.

Dentro de estos límites, el C.I. puede rendir y ha rendido servicios muy estimables como método de medida. Una de las clasificaciones más usadas de C.I. es la propuesta por Terman:

C.I.	Clasificación
Sobre 140	"Genio" o casi "genio"
120 - 140	Inteligencia muy superior
110 - 120	Inteligencia superior
90 - 100	Inteligencia media o normal
80 - 90	Torpe
70 - 80	Frontera con deficiencia mental
bajo 70	Deficiencia mental

El cuadro anterior debe ser interpretado con cautela. Solo podrá hacerlo una persona con suficiente formación psicológica. El término *genio*, por ejemplo, no tiene aquí significación corriente, sino otra más particular y limitada; indica simplemente una inteligencia general excepcional. Del mismo modo la denominación *deficiencia mental* se refiere únicamente a un nivel muy bajo de inteligencia y no equivale, necesariamente, al término jurídico que indica incapacidad de conducta normal o irresponsabilidad en los actos.

b) Normas centiles

Otra manera de establecer normas interpretativas consiste en ordenar las pun-

tuaciones del grupo normativo en una escala centil. Para ello es preciso averiguar los puntos centiles, o sea, las puntuaciones que dejan por debajo los sucesivos tantos por ciento, del uno al ciento por ciento, de los individuos de la muestra. Así, si solo el 3% de los sujetos de la muestra alcanzan o superan la puntuación del 123, 123 designará el centil 97, porque el 97% de los individuos no llegan a ella. Si, en adelante, un sujeto obtiene 123 puntos en el test, podemos afirmar que es superior al 97% de su población en el rasgo que mide el test. Para interpretar los centiles, hay que tener en cuenta, sin embargo, que son en realidad, números ordinales, y que la distancia entre un centil y el siguiente no es constante, sino mucho mayor en los extremos que en el centro de la escala.

c) Normas típicas

Para eliminar los inconvenientes que presentan las escalas anteriores, se ha recurrido al uso de normas y puntuaciones típicas. Una puntuación típica es igual a la diferencia entre la puntuación directa y la media del grupo, dividida por la desviación típica de la distribución. Siempre que las distribuciones sean aproximadamente iguales y, mejor, si son aproximadamente normales, las puntuaciones típicas son índices cuantitativos de significación universal. Si un sujeto obtiene en un test una puntuación típica igual a 1, sabemos que esto significa siempre que está situado una desviación típica por encima de la media de su grupo. Si la distribución es normal, esto nos indica, además, que el individuo es superior al 84% de los sujetos de su población. Las puntuaciones típicas se pueden convertir en una escala con el origen situado suficientemente por debajo de la media, por ejemplo, en el punto -5 sigmas, y con una unidad igual a una fracción de la

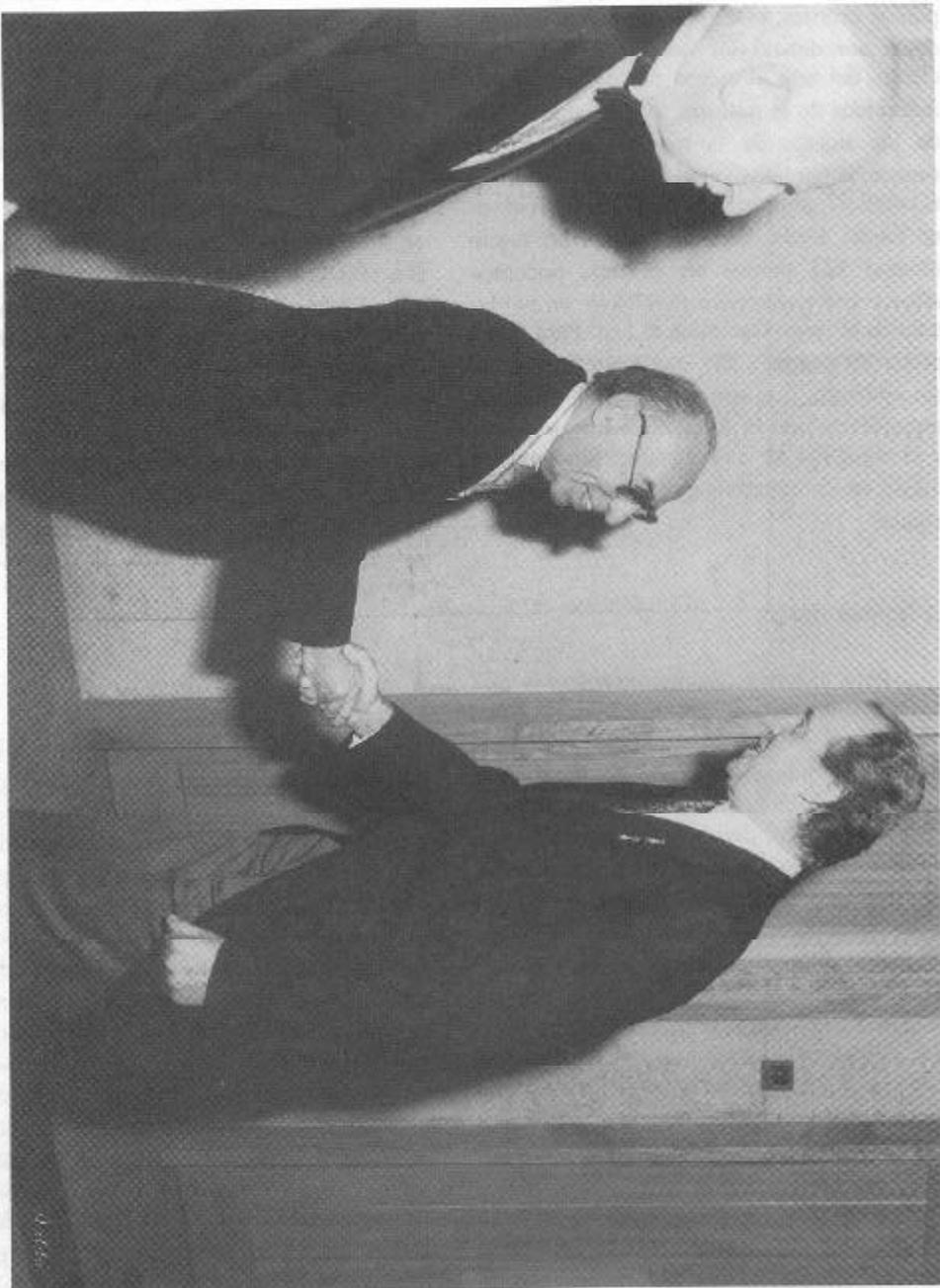
sigma, por ejemplo, igual a $0'10$ sigmas, con lo cual se tiene una escala de 0 a 100, que va desde -5 a 5 sigmas, de punto de partida fijo y de unidad constante.

Una de las escalas típicas más usadas es aquella en que el punto cero se sitúa en el punto $-2'5$ sigmas y la desviación típica se hace igual a 20 unidades. Con ello se tiene una escala que va de $-2'5$ sigmas a $+2'5$ sigmas, comprendiendo 100 unidades, desde 0 a 100, con la media en 50.

El único inconveniente de estas escalas es que su uso requiere algún conocimiento de estadística elemental. Por otra parte, esto puede considerarse como ventaja, ya que aparta así de los tests a los aficionados, cuya participación ha sido siempre deplorable para el prestigio del método y la orientación del examinador, y al mismo tiempo, exige que el psicólogo tenga, como debe, la competencia necesaria.

5. Aplicación de los tests

El test es un instrumento de medida y diagnóstico. Su eficacia depende tanto de su validez y precisión como de la competencia y pericia del que lo usa. El resultado de un test puede y debe expresarse mediante un número. Pero este número debe ser interpretado en relación con el comportamiento del sujeto durante el examen y en función de las circunstancias de su conducta. Todo test por exacto y válido que sea, refleja solo un aspecto parcial de la personalidad del sujeto en su momento, arbitrariamente fijado, de su vida. En el estudio del sujeto y de sus características psicológicas, el test es un medio, riguroso y fecundo, si se utiliza con discreción, pero no el único. Sus resultados deben interpretarse y ser complementados por otros métodos, a la luz de la historia personal del sujeto.



Real Academia. 1992. ENILLOS, YELA, REY D JUAN CARLOS