



## UN SIGLO UTILIZANDO TESTS.

**María José Navas**  
**Facultad de Psicología**  
**Universidad Nacional de Educación a Distancia**  
**e-mail: mjnavas@psi.uned.es**

### ABSTRACT.

The aim of this work is to think about tests in current society. In order to do that, some ideas related to the origin and development of tests are first highlighted. Then, issues regarding the usual areas of application of tests and the multiple functions they serve are dealt with. Next, there is a more detailed presentation of the more frequent critiques that tests have been suffering since the end of the 20s. The focus is on the heated debate on fairness and tests and on the misuse of tests. Finally, a few words are said about the future of tests and testing.

**Key words:** Test, origin, generalizability, critiques, future.

### 1. Introducción.

En la actualidad, los tests forman parte del entramado social en el que se desarrolla la vida del hombre; los tests acompañan al hombre durante una buena parte de su vida. Normalmente, desde los primeros años de escolarización, los niños comienzan a responder a tests que servirán para medir sus aptitudes intelectuales y su potencial para el aprendizaje; se utilizarán también para certificar su progreso y su nivel de conocimientos, así como para orientarlo vocacionalmente en el futuro. Más tarde, los tests pueden abrir (o cerrar) las puertas a la enseñanza universitaria y, una vez en el mercado laboral, los tests ayudarán a decidir quién deberá cubrir un determinado puesto de trabajo y quién debe ser promocionado; además, los tests desempeñarán también un papel importante a la hora de determinar si un sujeto está o no bien adaptado a su entorno, si dispone de recursos suficientes para resolver los problemas que se le plantean en su vida cotidiana o si, por el contrario, necesita algún tipo de terapia para reforzar alguno de estos aspectos.

El consumo de tests en nuestra sociedad es, por tanto, grande y grande es también el volumen de negocio que genera esta industria, sobre todo en el mercado anglosajón. Como dato ilustrativo, dos años y medio después de ser publicado el primer test de aplicación colectiva, ya habían sido evaluados con tests más de 4 millones de niños. A finales de los ochenta, más de un 90% de los centros y distritos escolares de EE.UU. utilizaban alguna forma de evaluación estandarizada; en 42 de los 50 estados la evaluación tenía lugar por orden

gubernativa y se estimaba que más de 100 millones de personas realizaban anualmente algún test estandarizado. En la actualidad, son muchos los millones de escolares, militares y opositores que son examinados regularmente con tests.

## **2.- El origen de los tests.**

El término test mental es acuñado por Cattell en 1890 y el primer test propiamente de inteligencia es creado a principios de siglo por los psicólogos franceses Binet y Simon. Este test surge para diagnosticar a niños con problemas de deficiencia mental en las escuelas públicas de París. Su acogida fue tan buena que poco antes del estallido de la I Guerra Mundial se da un reconocimiento institucional al papel de los tests en el procedimiento diagnóstico (Thompson y Sharp, 1988).

La I Guerra Mundial va a marcar un hito importante en la historia de los tests, al suponer la introducción por vez primera de tests de aplicación colectiva. Al entrar EE.UU. en la contienda, se ve en la necesidad de reclutar mucho personal de forma muy rápida, desarrollándose para ello los tests alfa y beta que fueron aplicados a cerca de dos millones de reclutas. Se trataba de dos tests muy sencillos, de aplicación colectiva y automática; el test alfa era una prueba verbal, adecuada para sujetos con capacidad lectora y el test beta una prueba no verbal, apropiada para sujetos con problemas de alfabetización.

El éxito de los tests en la selección y clasificación de los reclutas fue tal que rápidamente se generalizó su uso a la sociedad civil. Es decir, terminada la contienda, siguieron utilizándose dentro y fuera del ejército y su uso fue masivo.

La tecnología de los tests comienza a ser desde entonces una característica de la vida americana y no es en absoluto casual que sea justamente allí donde arraiguen y se desarrollen con fuerza los tests.

La importancia de los tests radica en que van a representar la posibilidad de juzgar a las personas por sus aptitudes, habilidades o conocimientos, esto es, por méritos propios y no por su nivel socio-económico, apariencia, o por el juicio subjetivo de profesores o supervisores. Parece claro que los tests se relacionan de forma más estrecha con la actuación posterior del sujeto en el puesto laboral o educativo o, en cualquier caso, los tests son dispositivos más objetivos que criterios como la cuna, la familia, la tribu, la apariencia y las opiniones intuitivas basadas en impresiones personales o en cartas de recomendación, criterios que han sido frecuentemente utilizados en situaciones de decisión.

Entonces, ¿qué mejor caldo de cultivo para los tests que EE.UU., la tierra de las oportunidades, tal y como ésta era percibida -o quizás soñada- por las sucesivas olas de inmigrantes que llegaban a EE.UU. en las primeras décadas de este siglo? Los tests fueron acogidos con los brazos abiertos por una sociedad meritocrática como la norteamericana, por una sociedad que creía en el mérito personal como forma de progresar y en la que cualquier persona que trabajara duro y con la suficiente capacidad podía llegar a la cima, podía tener éxito.

La sociedad norteamericana de comienzos de siglo constituye, por tanto, un terreno bien abonado para los tests, ya que era una sociedad bastante caótica, muy marcada por el fenómeno de la inmigración que hacía de ésta una población con un crecimiento muy rápido y con una enorme diversidad étnica y racial. Para poder organizar y educar a una sociedad tan diversa, era necesario poder detectar las diferencias individuales en las capacidades mentales de los sujetos (French y Hale, 1990). Si a esto se une la fuerte expansión de su economía y un enorme crecimiento industrial, tenemos ya los principales ingredientes del cóctel que explicaría el entusiasmo con el que fueron acogidos los tests en las primeras décadas del siglo en EE.UU. (Anastasi, 1985; Wigdor y Garner, 1982). De alguna manera, los tests sirvieron para clasificar y organizar en distintos ámbitos de la vida a las personas de esta sociedad en plena efervescencia durante la primera mitad del siglo.

### **3.- La generalización de los tests.**

Se produjo una rápida implantación de los tests en la industria y en las escuelas, no solo para el diagnóstico de niños deficientes sino también en la evaluación de niños normales, sin problemas. En 1920 aparece la primera batería estandarizada de rendimiento para evaluar distintas áreas académicas. Además, los tests podían servir para cubrir las necesidades de selección y de orientación individual que surgieron como consecuencia de la rápida expansión de la oferta educativa. Como ya se indicó, los tests acompañan a los escolares desde la educación infantil hasta el final de la enseñanza universitaria: los escolares se ven inmersos en un proceso más o menos continuo de evaluación diseñado para monitorizar su progreso, proporcionar una base para controlar el acceso a programas con plazas limitadas y para certificar el nivel de aprovechamiento y adquisición de conocimientos de cada alumno.

En la industria y en el medio empresarial en general, se espera que la evaluación de los sujetos y de las diferencias individuales proporcione un medio adecuado para seleccionar y clasificar a las personas para lograr el progreso económico y social. Se trata de seleccionar a la persona adecuada para el puesto adecuado y de minimizar, de este modo, los problemas con la mano de obra (Link, 1919). En este ámbito, los tests se usan fundamentalmente en la clasificación, selección y promoción laboral de los trabajadores. El uso de tests adecuados se ha demostrado que contribuye significativamente al aumento de la productividad y al tratamiento igualitario y justo de los sujetos en su entorno laboral. En 1919, Scott funda la Scott Company, primera consultora dedicada a la psicología industrial y empresarial, centrada en mejorar la eficiencia en la producción. En 1922 Cattell funda la Psychological Corporation, primera entidad dedicada a la publicación masiva de tests para la empresa, la industria y la educación. En la actualidad, esta compañía es no solo la más antigua de su sector sino también la de mayor peso, la editora de tests más importante en EE.UU. .

El tercer sector de aplicación de los tests es el clínico. En este ámbito, se suele realizar un uso diagnóstico de los tests, pero también se utilizan para predecir la conducta del sujeto y su respuesta al tratamiento o terapia, en el marco de la psicología de la salud. Asimismo, la evaluación clínica está ligada en ocasiones a decisiones que pueden tener un peso o un impacto muy importante y duradero en la vida de las personas. Así, la evaluación clínica se puede utilizar a la hora de decidir acerca de la veracidad de un testimonio, o para determinar la

custodia legal de los hijos en un proceso de separación y divorcio, para dictar una resolución temporal de libertad condicional o una sentencia firme, o para dilucidar acerca de la pertinencia de una alegación de perturbación mental transitoria a la hora de cometer el delito juzgado.

En resumen, las tres grandes áreas de aplicación de los tests serían la educativa, la industrial y la clínica, si bien existen también otros ámbitos de aplicación con ramificaciones en estos tres campos, como son la evaluación de programas, las certificaciones profesionales y académicas y el consejo u orientación.

En la evaluación de programas de cualquier tipo -educativo, sanitario, económico, de promoción de empleo, etc.- es muy frecuente el uso de los tests y, además, con múltiples funciones. En primer lugar, los tests se usan para vender el programa (cualquier intervención requiere una valoración previa y los tests, como instrumento de evaluación que son, tienen mucho que decir en este campo); los tests se usan también en la fase misma de implementación del programa y se utilizan posteriormente para determinar la bondad y eficiencia del mismo, esto es, sus resultados. Por ejemplo, en el ámbito educativo, los resultados de los tests proporcionan a las autoridades educativas información muy valiosa acerca del desarrollo académico general de los estudiantes y del funcionamiento general del sistema educativo, que les puede servir para reorientar su política de asignación de recursos, para modificar o redefinir objetivos curriculares.

El objetivo básico de los tests y pruebas utilizadas en la certificación profesional y académica es asegurar al público o a los potenciales usuarios de un profesional que éste es una persona cualificada y competente en su campo de especialización. Se trata de proteger a los usuarios, garantizando que las personas que han sido debidamente acreditadas profesional o académicamente poseen en grado suficiente los conocimientos y habilidades necesarias para desarrollar su actividad de forma apropiada. La certificación profesional y académica no es más que un mecanismo que sirve para identificar a los profesionales que satisfacen unos determinados criterios consensuados de competencia o excelencia.

Por último, el consejo u orientación puede ayudar a los sujetos a examinar distintas opciones educativas, familiares, relacionadas con la carrera profesional del sujeto, con su jubilación o simplemente con su ocio.

#### **4.- La polémica: tests sí, tests no.**

El abanico de campos de actuación y de fines cubiertos por los tests es enorme. Sin embargo, a pesar de ser muy utilizados han sido también -o quizás precisamente por eso- muy criticados. Muy pronto empieza a cambiar la consideración social de los tests, enfriándose algo el entusiasmo inicial que éstos habían suscitado; se comienza a escribir entonces el capítulo de críticas a los tests, capítulo todavía sin terminar al día de hoy. Desde la década de los 20 y hasta la fecha, los tests han sido criticados con más o menos virulencia y ferocidad, según las épocas, pero tanto las críticas como el uso de los tests ha sido una constante. Por tanto, después de un período inicial de uso acrítico y entusiasta de los tests, la tónica dominante ha sido un uso continuado y generalizado pero crítico. Quizás, si se comparan las primeras décadas del

siglo con esta última, el diagnóstico podría ser el siguiente: a las puertas del cambio de milenio, se dispone de más tests y de un mayor corpus teórico, pero se tiene menos confianza en ambos que a comienzos de siglo o, al menos, una confianza menos ciega.

Las voces críticas se empiezan a escuchar a finales de la década de los 20; en la década de los 40 la sociedad americana asiste ya a una versión preliminar de la moderna controversia sobre el cociente intelectual, los tests y la evaluación, que se desata con fuerza y alcanza su apogeo a finales de la década de los 60 y durante toda la década siguiente (Cole, 1986; Cronbach, 1975). A partir de los 80, baja un poco la temperatura del debate, pero el fuego todavía no se ha apagado.

Todos los tests estaban en el ojo del huracán pero, especialmente, los tests de cociente intelectual y de inteligencia o aptitudes mentales en general (Jensen, 1980). La realidad es que la evaluación psicológica realizada con tests no ha sido cuestionada cuando ésta se ha efectuado con un objetivo diagnóstico, habitualmente en el campo de la psicología clínica o de la salud; la evaluación psicológica con tests ha sido criticada, sobre todo, cuando los tests son utilizados en procesos de selección y, como no podía ser de otra forma, las voces más airadas proceden -lógicamente- de los que se quedan fuera en el proceso, de los que dejan fuera los tests en situaciones de selección en el ámbito laboral o educativo (Graham y Lilly, 1984).

Una crítica muy general que se les ha hecho a los tests es que no son ni lo suficientemente válidos ni lo suficientemente fiables como para justificar su uso. Los detractores más radicales sostienen que, aun en el mejor de los casos, los tests que realmente evalúan lo que pretenden lo hacen de una forma bastante pobre y, por tanto, no resulta adecuada su utilización en procesos de selección, orientación o distribución de recursos. Se cuestiona a los tests por medir, en ocasiones, habilidades muy específicas, de alcance muy limitado como para ser útiles para hacer predicciones significativas y a largo plazo. Es incuestionable, además, que los tests de inteligencia no miden muchas cosas que son importantes en la actuación posterior del sujeto en el puesto laboral o educativo, lo que limita parcialmente su capacidad predictiva (Wigdor y Garner, 1982).

Junto a estas críticas de carácter general, también se han cuestionado aspectos más puntuales de los tests, como el contenido y formato de sus ítems (especialmente, los de elección múltiple han sido objeto de frecuentes polémicas), las normas utilizadas para la interpretación de sus puntuaciones y la interpretación misma de las puntuaciones de los tests.

En ocasiones, los tests han tenido efectos secundarios en el proceso de enseñanza-aprendizaje de los escolares. En efecto, muchos profesores, ante las presiones sufridas dado el carácter obligatorio de la evaluación, se han limitado a enseñar las habilidades que posteriormente iban a ser evaluadas, pero es que, además, las han enseñado tal y como sugerían las especificaciones de los tests utilizados en cada momento y de ningún otro modo (teaching to the test), de manera que la instrucción se ha visto focalizada en exceso y los alumnos con dificultades para generalizar (Popham, 1992). Los estudiantes pueden dedicar gran parte de su tiempo a aprender cosas solo porque éstas van a ser posteriormente preguntadas en los tests utilizados en la evaluación y no porque tengan un valor intrínseco como objetivos educativos. Los tests estandarizados están siendo cada vez más criticados por

fracasar en proporcionar objetivos instruccionales adecuados.

Los efectos no esperados de la evaluación, cuando no son deseables, pueden ser seriamente perniciosos y distorsionar gravemente la esencia misma de la propia evaluación. Por consiguiente, el uso masivo -o, cuando menos, regular- de los tests en el ámbito educativo ha tenido de algún modo un efecto perverso en el sistema educativo ya que, en ocasiones, el curriculum se ha ido reorientando hacia los contenidos y habilidades evaluadas por los tests estandarizados, en lugar de ajustarse éstos a los objetivos curriculares definidos para los distintos cursos y ciclos educativos.

A los tests se les ha acusado también de ser herramientas al servicio del poder, de ser instrumentos de control o represión social al servicio de la clase económica y políticamente dominante, de constituir barreras para la igualdad social y de oportunidades económicas, simplemente porque los tests han servido para revelar diferencias entre grupos, y el tema de las diferencias entre grupos es muy espinoso, levanta muchas ampollas.

Los tests han sido puestos en la picota porque ciertos grupos sociales tienden a puntuar como grupo por debajo de la media de otros grupos socialmente más favorecidos. Por ejemplo, los norteamericanos negros puntúan como grupo una desviación típica por debajo del grupo de norteamericanos blancos en tests de inteligencia, es decir, hay una diferencia de 15 puntos en la escala de cociente intelectual entre la media de blancos y negros. Las mujeres como grupo también tienden a puntuar por debajo de los hombres en algunas pruebas de habilidad espacial y de razonamiento matemático, y por encima de los hombres en pruebas de aptitud verbal.

El problema es que, durante bastante tiempo, las diferencias observadas en los tests entre distintos grupos sociales, étnicos o raciales han sido atribuidas a los tests, es decir, han sido explicadas aduciendo que los tests estaban sesgados contra distintas minorías o grupos, favoreciendo sistemáticamente a unos grupos en perjuicio de otros. Esto tiene una repercusión social muy importante: el uso de los tests en procesos de selección conduce a tasas distintas de selección en los distintos grupos. La preocupación básica es dictaminar si los tests utilizados habitualmente en el ámbito educativo y empresarial -y contruidos, normalmente, por la clase y raza económica y políticamente dominante- están sesgados contra grupos minoritarios, es decir, si conducen a una discriminación en el acceso al mercado laboral y educativo de personas pertenecientes a clases y razas ni económica ni políticamente dominantes.

Las denuncias sobre la parcialidad de los tests crecen considerablemente a finales de la década de los 60 y, sobre todo, durante los años 70, en parte, debido al fuerte impulso que experimentó en esa época el movimiento por los derechos humanos, que va a desatar un enorme interés y preocupación por el posible uso y efecto discriminador de los tests.

Los tests pasan a disposición judicial y su adecuación en distintos contextos de aplicación se va a dirimir en los juzgados. De hecho, en algunos estados americanos se prohibió el uso de tests en la toma de decisiones educativas, se fallaron sentencias judiciales importantes por discriminación en la selección de personal y en la admisión a instituciones educativas y se cometieron también no pocas tropelías desde el punto de vista técnico. Una de

las más conocidas es el caso Golden Rule. Tras ocho años de litigios, el Educational Testing Service y la compañía de seguros Golden Rule acordaron eliminar de las pruebas que utilizaban con sus agentes de seguros aquellos elementos que eran respondidos correctamente un 15 % más en la población blanca que en la afroamericana. Esta decisión, políticamente rentable, era técnicamente un disparate.

Para acabar de complicar las cosas, la polémica sobre los tests se vio metida de lleno en el ya de por sí controvertido debate sobre la herencia y el medio (véase, por ejemplo, Block y Dworkin, 1976; Jensen, 1969, 1980 y Reynolds y Brown, 1984), cargando el acento más en el lado hereditario que en el ambiental. Eran muchos -aunque no todos- los autores que sostenían que las diferencias en las puntuaciones de los tests de distintos grupos reflejaban diferencias hereditarias en habilidad, y que los tests de inteligencia general medían características genéticas.

Así las cosas, es fácil entender porqué los tests eran considerados como instrumentos claramente reaccionarios, ya que podían proporcionar un fundamento a la desigualdad social (Gould, 1981), al estar ésta basada o justificada por las diferencias en aptitud de distintos grupos.

Esta situación resulta, si no paradójica, sí cuando menos curiosa. En efecto, los tests que fueron acogidos con gran entusiasmo por suponer un paso adelante frente a posiciones retrógradas que utilizaban como criterios para medrar la posición social, la familia o la cuna, pocas décadas después son vilipendiados justamente por grupos o minorías socialmente desfavorecidas. Como señala Cronbach (1975), lo llamativo es que los impulsores de los tests pretendían abrir con ellos las puertas a personas competentes pero socialmente más desfavorecidas e, irónicamente, el ataque más feroz procede justamente de los que hablan en nombre de los pobres.

Lo cierto es que la preocupación por el tratamiento igualitario de grupos y minorías, aunque totalmente legítima, condujo a una situación equívoca, llevó a un planteamiento simplista y algo ingenuo conocido como la falacia igualitarista.

Esta falacia arranca del supuesto de que todos los hombres son iguales, sin distinción alguna de raza, color, sexo, religión, origen social o nacional. Este supuesto es totalmente gratuito, ya que a nadie se le escapa la diferencia que existe entre nacer en un país como España y, por ejemplo, en Somalia; la diferencia es enorme no ya en las oportunidades sociales y económicas sino simplemente en términos de mera supervivencia. Pues bien, partiendo de este supuesto -falso- de que todos los hombres son iguales, se plantea que si los tests hacen a los hombres desiguales es porque los tests son injustos, ya que la justicia está obviamente con la igualdad de oportunidades. Así, con esta argumentación, cualquier test en el que se encontrasen diferencias entre grupos étnicos, culturales o socioeconómicos, era considerado injusto y sesgado.

La realidad es que la falacia igualitarista llevó a confundir la legítima igualdad de derechos y oportunidades con la (des)igualdad de los resultados obtenidos. Esto condujo, irremisiblemente, a matar al mensajero -los tests- cuando éstos lo único que hacían era

constatar -que no producir- las diferencias en la actuación media de distintos grupos. Los tests eran solo el mensajero, no el enemigo: se confundían las diferencias entre grupos con el sesgo y se trataba ingenuamente de buscar la igualdad en los resultados como forma de remediar la desigualdad de oportunidades. Sin embargo, el papel de los tests no es conjurar los demonios sociales del mundo sino ser un reportero neutral de lo que en el mundo acontece.

Trasladado el problema a la arena política, su solución pasa por encontrar un punto de equilibrio entre el principio de igualdad de oportunidades para cada individuo de la sociedad con la realidad incontestable de una procedencia y una formación desigual de dichos individuos.

Trasladado el problema a la arena técnica, el objetivo de los constructores de tests y de los psicómetras debe centrarse en las diferencias entre grupos que tienen su origen en los ítems del test, y no en el ambiente o la procedencia de los individuos. Esos ítems son los que realmente pueden sesgar o distorsionar la actuación de los sujetos en el test.

En suma, los tests han sido criticados o debatidos tanto dentro como fuera del marco de la psicología, es decir, las críticas han trascendido el ámbito de lo puramente psicométrico y los tests se han convertido en una cuestión de debate público e incluso legal.

En ambos terrenos -público y profesional- las críticas a los tests tienen que ver más con el uso inadecuado que a veces se hace de los tests que con las propias características o propiedades técnicas de los mismos. Para ilustrar este punto, Muñiz (1998) compara los tests con el automóvil, considerado como el mayor homicida de nuestros días. Sin embargo, a pesar del elevado número de muertes que ocasiona, son muy pocos los accidentes atribuibles a fallos mecánicos, es el uso inapropiado que se hace del automóvil lo que genera la elevada tasa de siniestralidad. Algo similar ocurre con los tests: el problema no es tanto sus características técnicas sino el uso que se hace de ellos. Y el mal uso de los tests puede provenir de tres fuentes: la negligencia en su uso, la mala fe, es decir, un intento deliberado de distorsionar la realidad y conocimientos insuficientes, falta de información o información errónea. Esta última fuente es, con diferencia, la causa más frecuente del uso inapropiado de los tests.

Atajar este problema supone necesariamente regular el uso de los tests. La cuestión es cómo y cuánto.

Las fuentes de regulación del uso de los tests pueden proceder desde distintos frentes. Por ejemplo, se puede trabajar desde la propia industria de los tests, desde organizaciones o colegios profesionales o desde la misma administración pública (Wigdor y Garner, 1982). En cualquier caso, se trata de maximizar la excelencia técnica de las pruebas y la preparación de los usuarios. No hay recetas mágicas para ello, pero parece que la mejor fórmula consiste en combinar la formación de los usuarios con la restricción en el acceso a los tests de personas no cualificadas aunque, a la larga, los esfuerzos por mejorar la formación son más eficientes que la mera restricción en el uso (Fremer, 1996; Tyler, 1986).

Desde hace tiempo, el Colegio Oficial de Psicólogos viene realizando diversas actividades para mejorar el uso y la práctica de los tests en España.

A nivel nacional, se formó en su día una comisión de tests constituida por universitarios, profesionales y casas editoras que trata de discutir y proponer medidas tendentes a la mejora del uso de los tests en España (véase la página web del colegio <http://www.cop.es> para obtener información relativa a estas medidas).

A nivel internacional, el Colegio Oficial de Psicólogos participa en la International Test Commission, organismo encargado de potenciar un uso adecuado de los tests a nivel internacional. Este organismo está ultimando unas normas internacionales para el uso de los tests que verán la luz en el 2000. Estas normas se articulan en torno a dos grandes apartados: el uso ético de los tests y la utilización adecuada de los tests.

Existe también una comisión europea para el estudio y mejora de la construcción y uso de los tests en Europa.

## **5.- A las puertas del nuevo milenio.**

La realidad es que los tests se siguen utilizando en todos los ámbitos de aplicación en los que tradicionalmente han sido empleados y con las mismas funciones que habitualmente cubrían. Eso sí, sin el aura de autoridad y de validez con la que parecían investidos durante la primera mitad de siglo.

La fuerte polémica suscitada por los tests ha servido un poco para volver las aguas a su cauce. Así, al mostrar abiertamente las limitaciones de los tests, se ha neutralizado la tendencia inicial a considerarlos como la panacea a un buen número de problemas o situaciones. Del mismo modo que la discusión social que han generado los tests ha servido también finalmente para contrarrestar la tendencia -igualmente prevalente- a tomar a los tests como cabeza de turco o chivo expiatorio de problemas o guerras que, en realidad, no eran las suyas.

A pesar de todas sus limitaciones e inconvenientes, todavía no se ha propuesto un método mejor, o más objetivo que los tests. Cuando lo que se pretende es tomar decisiones acerca de un sujeto y en esta decisión está implicado su funcionamiento intelectual, los tests siguen siendo mejores herramientas diagnósticas que criterios como, por ejemplo, la intuición del psicólogo o las cartas de recomendación (Fernández Ballesteros, 1993). Ahora bien, hay una serie de consideraciones que siempre hay que tener en cuenta.

Primero, los tests no deben ser la única fuente de información a la hora de tomar una decisión importante para un sujeto, no se pueden utilizar como único criterio en la toma de decisión sino que es importante acudir a otras fuentes.

Segundo, la puntuación en un test es siempre un indicador solo imperfecto de la variable psicológica que mide el test. Además, un test puede ser un buen predictor de la conducta o actuación posterior de los sujetos en general, como grupo, pero puede resultar ser un predictor pobre del comportamiento de un sujeto particular.

En la actualidad, hay elementos que permiten apostar con cierto optimismo por el futuro de los tests. Si bien es cierto que el uso de los tests ha sido bastante controvertido, no lo es menos que nunca se han dejado de utilizar. Se han aprobado leyes que prohibían o limitaban su uso, pero también se han aprobado leyes que fomentaban u obligaban a utilizar tests. Buena prueba de ello son, por ejemplo, los programas promovidos por muchos estados americanos de evaluación de competencias escolares mínimas y el programa NAEP, programa creado por mandato del congreso estadounidense cuyo objetivo es evaluar de forma permanente el sistema educativo americano. Este tipo de programa es cada vez más frecuente en un buen número de países, y no solo a escala nacional sino también regional y local, e incluso a nivel internacional. Estos programas obedecen a la necesidad cada vez más imperiosa de responder o rendir cuentas del dinero invertido en educación por las administraciones regionales y nacionales.

En el futuro, la práctica profesional con tests deberá ser extremadamente cauta y prudente y tendrá que caracterizarse por una mayor transparencia en sus métodos y formas de operar. Ésta es, sin duda, una buena vía para acallar críticas y evitar procesos judiciales. No hace muchos años, en EE.UU. se consideraba innecesario informar al sujeto de su puntuación en una prueba de acceso a la Universidad. En la actualidad, no solo es preciso informarle de su puntuación sino que es obligatorio mostrar públicamente el test y su plantilla de corrección. Éste es el camino por el que parece que avanza el futuro de los tests.

Los tests del futuro apuntan hacia el diseño de formas menos restrictivas de ítems y hacia la construcción de nuevas formas de tests, tanto en su contenido como en su formato y modo de administración.

Hacia el final de la década de los ochenta se comienza a ampliar el curriculum para incorporar habilidades de pensamiento y razonamiento de orden superior. Las habilidades integradoras, metacognitivas y las estrategias de solución de problemas se consideran importantes. En esta nueva perspectiva, el foco de atención de investigadores, educadores y legisladores se va desplazando hacia la así llamada medición auténtica (authentic measurement) o evaluación de la actuación (performance assessment). Las exigencias crecientes de evaluar no solo habilidades básicas sino también habilidades cognitivas de orden superior suponen un nuevo reto para los tests, ya que implican el recurso a pruebas con ítems de respuesta abierta y, especialmente, pruebas de respuesta construida que pueden proporcionar un conocimiento más profundo y un análisis más detallado de niveles superiores en la actuación de los sujetos que las clásicas pruebas de elección múltiple. El abanico de formatos posibles es enorme, desde pequeñas adaptaciones de ítems de elección múltiple hasta tareas del tipo de realizar informes orales o escritos, presentaciones, revisiones, proyectos, etc., en suma, tareas muy semejantes a las actividades reales de aprendizaje del alumno.

A todo esto, hay que añadir el tremendo impulso que la tecnología del ordenador ha proporcionado al campo de los tests, posibilitando su administración a través del ordenador y la construcción de bancos de ítems y de tests adaptativos informatizados.

El camino recorrido por los tests tras un siglo de andadura ha sido espinoso pero fecundo. Hay que convenir con Muñiz (1998) en que los tests constituyen probablemente la

tecnología con mayor impacto social que ha producido la psicología. Ahora bien, el futuro dista mucho de ser un camino de rosas. Son muchos los desafíos que se plantean a los tests para responder a las crecientes demandas -cada vez más exigentes y diversificadas- que se generan en la sociedad y a la necesidad de una mayor claridad y transparencia a lo largo de todas las fases implicadas en la utilización de los tests, desde su construcción hasta el preciso momento de informar a los sujetos acerca de las puntuaciones obtenidas en los mismos.

## 6.- Referencias.

- Anastasi, A. (1985). Some emerging trends in psychological measurements: A fifty year perspective. *Applied Psychological Measurement*, 9(2), 121-138.
- Block, N. J. y Dworkin, G. (Eds.) (1976). *The IQ controversy*. Nueva York: Pantheon.
- Cole, N. S. (1986). Future directions for educational achievement and ability testing. En B. S. Plake y J. C. Witt (Eds.), *The future of testing. Buros-Nebraska: Symposium on Measurement and Testing*.
- Cronbach, L.J. (1975). Five decades of public controversy over mental testing. *American Psychologist*, 30, 1-14.
- Fernández-Ballesteros, R. (1993). Evaluación psicológica en sus contextos de aplicación. *Revista de Historia de la Psicología*, 14, 97-117.
- Fremer, J. (1996). Promoting high standards for test use: Developments in the United States. *European Journal of Psychological Assessment*, 12, 160-168.
- French, J. L. y Hale, R. L. (1990). A history of the development of psychological and educational testing. En C. R. Reynolds y R. W. Kamphaus (Eds.), *Perspectives on bias in mental testing*. Nueva York: Plenum Press.
- Gould, S. J. (1981). *The mismeasure of man*. Nueva York: Norton.
- Graham, J. y Lilly, R. S. (1984). *Psychological testing*. Englewood Cliffs, NJ: Prentice-Hall.
- Jensen, A. R. (1969). How much can be boost IQ and scholastic achievement? *Harvard Educational Review*, 39, 1-123.
- Jensen, A. R. (1980). *Bias in mental testing*. Nueva York: The Free Press.
- Link, H. C. (1919). *Employment psychology: The application of scientific methods to the selection, training, and grading of employees*. Nueva York: Macmillan.

- Muñiz, J. (1998). *Teoría clásica de los tests*. Madrid: Pirámide.
- Popham, W.J. (1992). A tale of two test specification strategies. *Educational Measurement: Issues and Practice*, 11(2), 16-17, 22.
- Reynolds, C. R. Y Brown, R.T. (Eds.) (1984). *Perspectives on bias in mental testing*. Nueva York: Plenum Press.
- Thompson, G. O. B. y Sharp, S. (1988). History of mental testing. En J. P.Keeves (Ed.), *Educational research, methodology and measurement. An international handbook*. Oxford: Pergamon Press.
- Tyler, B. (1986). Responsibility in practice: Some implications of the BPS survey on the test use. *Bulletin of the British Psychological Society*, 39, 410-413.
- Wigdor, A.K. y Garner, R. (Eds.) (1982). *Ability testing: Uses, consequences and controversies*. Washington, DC: National Academy Press.