



**Nombre de la alumna: Fany Giselle Oseguera Fino.**

**Nombre del profesor: Luis Miguel.**

**Nombre del trabajo: Resumen.**

**Materia: Psicometría.**

**Grado: 3°**

**Grupo: Psicología.**



## **APROXIMACIÓN HISTÓRICA Y CONCEPTOS BÁSICOS DE LA PSICOMETRÍA.**

### **✚ La psicometría en el contexto de la psicología.**

Una aproximación histórica a la psicometría debe empezar recogiendo lo que se consideran sus orígenes y antecedentes. A pesar de que el desarrollo y la administración de test psicológicos es una práctica desarrollada fundamentalmente a partir del siglo XX, es posible encontrar algunos antecedentes remotos en culturas tan antiguas como la china. Partiendo de estos antecedentes, y repasando algunas de las contribuciones más importantes que han contribuido a su desarrollo como disciplina científica, estaremos en disposición de ofrecer una definición formal de la psicometría que nos permita situarla en el contexto general de la psicología.

### **✚ Una aproximación histórica a la psicometría.**

Como antecedentes remotos de la psicometría, tanto la evaluación en la educación formal como el establecimiento del sistema de evaluación imperial supusieron un importante cambio en la concepción sobre el juicio de las capacidades de las personas. De este modo, se fue trasladando la confianza en el juicio personal basado en impresiones, hacia la administración de pruebas institucionales basadas en una autoridad imparcial, que objetivaba las destrezas y los conocimientos requeridos en los ámbitos educativo y administrativo. De acuerdo con el trabajo de Bondy (1974), es conveniente tener presente el esfuerzo de los primeros profesionales orientados al estudio de los problemas mentales y las lesiones cerebrales en el establecimiento de lo que se podrían considerar las primeras pruebas de evaluación psicológica. Así, por ejemplo, se desarrollaron los primeros tests para evaluar las consecuencias del daño cerebral, alguno de ellos tan elaborado que exigía una administración durante periodos de 100 horas. A pesar de sus limitaciones, entre otras la ausencia de procedimientos estándares en su uso, estas primeras pruebas incorporaron muchos de los elementos que, con la evolución del estudio de los síntomas psicológicos, todavía son utilizados en las pruebas de diagnóstico actuales.

Los antecedentes recientes de la psicometría se encontrarían en el desarrollo del estudio sistemático de las diferencias humanas durante el siglo XIX. Primero gracias a los trabajos de Friedrich W. Bessel (1784-1846) y Carl F. Gauss (1777-1855) –que fueron pioneros en el estudio de las diferencias individuales en la percepción en el campo de la astronomía– y después a las contribuciones de Gustav T. Fechner (1801-1887) y Hermann von Helmholtz (1821-1894) en el desarrollo de la psicofísica –que supuso el inicio de la psicología como disciplina académica–, ambas aproximaciones constituyeron un avance importante en la sistematización de la medida de las sensaciones psicológicas producidas por la estimulación física (Boring, 1978, para una revisión histórica).

Así, mediante el desarrollo de las primeras leyes que relacionan estímulos físicos y sensaciones, se fueron asentando las bases para la medida psicológica.

De este modo se perfeccionaron los métodos de presentación de los estímulos y el registro de las respuestas, se trabajó en la mejora de la precisión de las medidas y se adoptaron condiciones controladas para su consecución. Todos estos elementos son indispensables para el planteamiento posterior de los problemas más importantes que la psicometría debería afrontar en su desarrollo inicial.

Gracias a la adopción de un grupo de referencia compuesto por niños de entre 3 y 13 años, y una vez ordenados los ítems en función de la edad en la que eran típicamente resueltos, desarrollaron el primer test para cuantificar la inteligencia. Esta cuantificación, tomando como referencia la edad biológica del niño, condujo a la definición de lo que se conoce, hasta la actualidad, como edad mental.

El prolífico Thurstone, que desarrolló las primeras pruebas de análisis de consistencia interna de los tests en el ámbito de la evaluación de la personalidad, hizo otra contribución importante al desarrollo de la psicometría. Partiendo de los experimentos desarrollados por los pioneros (Gulliksen, 1968), especialmente en la rama de la psicofísica, propuso en 1927 la ley del juicio comparativo como el método para la medida de las actitudes, las preferencias y los valores. Esta innovación consistió en trasladar el juicio perceptivo sobre parejas de estímulos físicos –por ejemplo, la elección entre dos sonidos según su intensidad– a la valoración de características psicológicas no estrictamente relacionadas con las propiedades físicas –por ejemplo, la elección entre dos comportamientos según su aceptabilidad. Después de proponer una segunda ley, la ley del juicio categórico, la aproximación de Thurstone a la medida de las actitudes estaba preparada para ser aplicada al juicio general, es decir, sin requerir la comparación entre parejas de fenómenos. De este modo, tal y como pretendían Thurstone y sus continuadores (Bock y Jones, 1968), la psicometría dispondría de un método de escalamiento para medir y tratar numéricamente valoraciones individuales subjetivas en la búsqueda de una representación objetiva de los fenómenos psicológicos. Este método dio un soporte metodológico importante en las décadas siguientes de investigación en psicología social.

### **La psicometría hoy.**

La psicometría, un nuevo espacio de trabajo metodológico en torno al desarrollo y a la administración de test, se fue conformando con un componente aplicado, orientada a las demandas en diferentes contextos, y formó parte de algunos de los debates teóricos más importantes de la propia psicología.

El periodo entre los años treinta y sesenta como la época dorada de la psicometría. En este periodo es cuando se publican, además, los libros y manuales más importantes que la vertebrarían. Entre los manuales hay que destacar *The reliability and validity of tests*, de Thurstone (1931), que sistematizaba lo que se había desarrollado hasta el momento en relación con la teoría de los tests y sugería el papel central de la fiabilidad

como requisito para la validez de las medidas en la psicometría. Volveremos más adelante a estas cuestiones, elementos básicos de la medida indirecta de los fenómenos psicológicos mediante tests. De manera contemporánea, fue también importante la primera edición del manual *Psychometric methods* de Guilford (1936), un intento de organizar el campo propio de la psicometría en torno a la teoría de los tests, el escalamiento psicológico y el psicofísico. La teoría clásica de los tests, como fue ampliamente conocida a partir de los trabajos de Spearman sobre la estimación de los errores de medida, empezaba su camino en los circuitos docentes universitarios encargados de formar a los futuros psicólogos.

Stevens definió la medida como el proceso de asignación de números a objetos o acontecimientos de acuerdo con unas reglas, producto de las cuales se obtendrían los diferentes tipos de escalas propuestas: nominal, ordinal, de intervalo y de razón. Medir, en este sentido, no consistiría únicamente en cuantificar, sino que sería el producto de la utilización de diferentes reglas que, en último término, determinarían el tipo de operaciones –o técnicas– estadísticas permitidas en cada escala. No exenta de críticas, es importante resaltar la importancia de su contribución como un primer intento para superar las restricciones impuestas por la cuantificación de los atributos físicos con el objetivo de resolver la controversia en torno a la medida de los fenómenos psicológicos. Así, el debate sobre los modelos o paradigmas de medida formó parte importante de la agenda psicométrica y se pusieron las bases para el desarrollo posterior de otros modelos, como el operacional y el clásico, sobre los que volveremos más adelante. Por otro lado, cerrando este breve repaso a las contribuciones importantes durante los años cuarenta, no podemos dejar de hacer mención a la publicación por parte de Thurstone (1947) del influyente *Multiple factor analysis*, que a partir de los trabajos de Spearman, Kelley y Burt proporciona el soporte estadístico necesario para la construcción y validación de los tests durante las siguientes décadas.

### **La psicometría en el contexto de la evaluación psicológica.**

Un primer grupo de definiciones aborda la psicometría a partir de los instrumentos que utiliza. Desde esta corriente se circunscribe la psicometría como la disciplina encargada de desarrollar los fundamentos para la construcción y administración de tests (por ejemplo, ver Martínez Arias, Hernández Lloreda y Hernández Lloreda, 2006). Esta definición, a pesar de que corta y concisa, presenta algunas dificultades. En primer lugar, la palabra test es polisémica, por lo que según el contexto en el que sea empleada puede ser sinónimo de otros términos con significados diferentes, como prueba, examen o incluso ensayo. En segundo lugar, más allá de la precisión semántica que se puede sobreentender en el contexto de la psicología, no deja de ser en cierto modo restrictiva. Si bien es cierto que los tests psicológicos son los instrumentos específicos que desarrollan y administran los psicómetros, también lo es que son el resultado aplicado de un proceso más amplio caracterizado por el desarrollo de teorías y métodos orientados a la medida indirecta de los fenómenos psicológicos. Una segunda corriente define la psicometría a partir del objeto al que somete o aplica su interés. Así, la psicometría se define de manera más o menos genérica como la

disciplina científica orientada a la evaluación o medida de los fenómenos psicológicos (por ejemplo, ver Rust y Golombok, 2009). Esta definición se ajusta de hecho al significado etimológico de la palabra psicometría, que en sus orígenes griegos podemos encontrar en la yuxtaposición de las palabras psique –que significa ‘alma’, ‘aliento’ o ‘intelecto’– y metron –en referencia al proceso de medida. En esta línea se puede situar la definición de Kline (1998), que aborda específicamente la psicometría como la tarea de desarrollar medidas científicas fundamentales en las áreas de la personalidad y las capacidades. En este contexto, la científicidad se convierte en sinónimo de estandarización y, a pesar de no hacer referencia explícita, alude indirectamente al uso de tests en el contexto del método científico. Otro implícito importante es, con relación al estatus métrico de las puntuaciones obtenidas como resultado del proceso de medida psicológica, el carácter cuantitativo que debería tener la psicometría como disciplina científica.

La psicometría, en este sentido, contribuye al desarrollo de la evaluación psicológica proporcionando teorías, métodos y técnicas que, en última instancia, permiten describir, clasificar, diagnosticar, explicar o predecir los fenómenos psicológicos objeto de medida. De hecho, contextualizada así la psicometría podemos ir un paso más allá y señalar que, a su vez, la evaluación psicológica no se encuentra en ningún otro lugar que al servicio de la intervención psicológica. Es decir, la finalidad última que guía el desarrollo y la administración de tests es la de contribuir a la recogida de las evidencias necesarias que permitan a los psicólogos tomar una decisión u orientar alguna acción. Evidentemente, esta intervención dependerá del contexto en el que la evaluación haya sido desarrollada, pero es importante tenerlo presente para entender no solo qué es la psicometría, sino también su importancia para el ejercicio profesional de los psicólogos en los diferentes contextos en los que intervienen.

### **Definición y clasificación de los tests.**

Un test psicológico es un procedimiento de medida objetiva y estandarizada de una muestra de comportamientos. Otras definiciones son posibles<sup>4</sup> pero esta contiene tres elementos fundamentales que nos permiten abordar sistemáticamente las características más importantes que cumplen de manera genérica los tests. En primer lugar, la medida mediante el desarrollo y la administración de tests es, o al menos pretende ser, objetiva. En este sentido, la objetividad hace referencia a la sustitución del juicio personal basado en criterios subjetivos por un conjunto de normas determinadas y conocidas que permiten obtener e interpretar las puntuaciones de los individuos en igualdad de condiciones. Asimismo, la medida que pretenden los tests es estandarizada, en cuanto que las puntuaciones obtenidas dependen de un procedimiento establecido de administración, corrección e interpretación que las hacen, o las deberían hacer, invariantes del profesional que administra los tests, las condiciones específicas en las que lo hace y el modo como obtiene e interpreta las puntuaciones resultantes de la media. Finalmente, los tests se enfrentan a los fenómenos psicológicos no observables mediante una muestra de comportamientos. En la medida en que esta muestra sea representativa del conjunto, las puntuaciones

obtenidas permitirán a los profesionales establecer adecuadamente sus inferencias sobre el comportamiento general de las personas más allá de los elementos específicos evaluados mediante test.

El objetivo habitual de la medida de las magnitudes físicas es obtener información sobre un único objeto. La psicometría, en cambio, se propone desarrollar instrumentos que permitan obtenerla sobre un grupo de individuos con el objetivo de extraer conclusiones sobre cada uno de ellos, sobre el grupo entero, e incluso extrapolar sus resultados a las poblaciones de referencia de donde provienen estos individuos. Por otro lado, la medida de magnitudes físicas parte de la posibilidad de repetir su procedimiento tantas veces como sea deseado sin variar las condiciones en las que se lleva a cabo la medida. Este procedimiento es incompatible con la medida mediante tests, en cuanto que la repetición en la aplicación de una misma prueba sobre los mismos individuos produce variaciones en las puntuaciones, que pueden ser explicadas por el cansancio o por el efecto de la práctica en el aprendizaje y no por variaciones sustanciales en los fenómenos de interés.

Estamos en disposición de clasificarlos en función de algunas de sus características más importantes:

- Según el propósito:

Los tests pueden tener diferentes finalidades, y entre ellas podemos destacar dos fundamentales: la diagnosis, orientada a la evaluación de las condiciones actuales de los individuos, y la predicción, que se propone relacionar la medida actual con el comportamiento de las personas en situaciones futuras. En este sentido, por ejemplo, una cosa es obtener información sobre la destreza de una persona en la resolución de problemas y otra es utilizar esta información para predecir su desempeño en un trabajo determinado.

- Según el contenido:

De manera general, podemos clasificar los tests según el área de la psicología a la que pertenecen los fenómenos psicológicos que pretenden medir. Esta clasificación, sin embargo, varía según los autores que las realizan, en cuanto que pueden fijarse en diferentes niveles de complejidad al jerarquizar estas áreas. En el nivel más general podemos distinguir básicamente tres grandes grupos, que incluirían los tests orientados a la evaluación de las habilidades cognitivas, la personalidad y las actitudes.

- Según el formato:

Teniendo en cuenta los materiales utilizados es frecuente distinguir los tests de lápiz y papel, de manipulación y de medidas fisiológicas. Nos referimos a los tests de lápiz y papel cuando presentan los ítems o preguntas en papel y requieren que el individuo dé algún tipo de respuesta escrita. Los tests computerizados, a pesar de que no usan el papel, podrían ser incluidos en esta categoría en cuanto que exigen algún tipo de

respuesta escrita mediante dispositivos electrónicos. Por otro lado, son tests de manipulación aquellos que presentan una serie de objetos, imágenes o rompecabezas que los individuos deben resolver para demostrar su habilidad. Finalmente, los tests de medidas fisiológicas utilizan sensores de distinto tipo para registrar las reacciones de los individuos ante los estímulos físicos presentados.

- Según el tipo de administración:

Otra manera de clasificar los tests puede tener en cuenta el modo como son administrados y nos permite distinguir los tests individuales y grupales. Esta clasificación se solapa con la que hemos hecho según el formato, pero nos permite distinguir tests que requieren una administración a un único individuo o permiten hacerlo con un grupo de individuos al mismo tiempo.

- Según el tratamiento de las respuestas:

Podemos clasificar también los tests en dos grandes grupos, en función de si evalúan el acierto en las respuestas para determinar la competencia, la aptitud o los conocimientos de los individuos, o si tratan de evaluar motivaciones, preferencias, opiniones o actitudes. Estos dos tipos de pruebas son también conocidas como pruebas de ejecución máxima y pruebas de ejecución típica, respectivamente.

- Según la interpretación de las puntuaciones:

Las puntuaciones obtenidas mediante tests pueden ser interpretadas según la norma o en referencia a un criterio. Así, los tests normativos permiten comparar la puntuación del individuo con la ejecución observada en un grupo de referencia que previamente ha respondido al mismo test. Este grupo puede estar compuesto por muestreo probabilístico, cuando representa a la población de referencia, o no probabilístico, cuando estamos interesados en comparar las puntuaciones con un grupo de personas que cumple unas características determinadas. Por otro lado, los tests referidos a criterio toman como referencia la definición de un dominio de conocimientos o habilidades específicos y permiten medir la ejecución del individuo no ya en comparación con un grupo de referencia, sino en función de su grado de adecuación o consecución de este criterio.

- Según el estatus comercial:

Los tests se pueden clasificar en función de si son propietarios o no. Los tests propietarios o comerciales son tests que requieren el pago para su uso e incluyen los ejemplares del propio test y un manual que contiene información sobre su desarrollo y sus propiedades psicométricas, las hojas de corrección de las respuestas y las tablas para interpretar las puntuaciones en relación con los grupos de referencia. En otros casos hablamos de tests abiertos o no comerciales, y generalmente se pueden obtener contactando con el autor, que normalmente también ha desarrollado algún tipo de manual que contextualiza el test.

## Modelos de medida psicométrica.

Son los modelos o paradigmas de medida que la psicometría ha desarrollado de manera formal a partir del debate iniciado en los años cuarenta sobre la viabilidad de la medida de los fenómenos psicológicos. No obstante, antes de caracterizar estos modelos retomaremos la controversia a partir de las conclusiones de Ferguson y sus colaboradores (1940) para la British Association for the Advancement of Science.

Campbell (1928) extendió este modelo general a lo que denominó como medidas derivadas, donde la cuantificación de ciertos atributos físicos no dependería de la manipulación empírica sino del descubrimiento de relaciones matemáticas entre dos medidas fundamentales. Por ejemplo, sería el caso de la densidad de los objetos, medida derivada que respondería a una razón entre su masa y su volumen. En cualquier caso, la clave de la medida para el modelo de Campbell se encontraría en la necesidad de observar y manipular empíricamente los objetos, bien para cuantificar un atributo físico mediante su observación directa, bien para establecer funciones matemáticas que relacionen otros atributos observables. En la medida en que los fenómenos psicológicos no pueden ser observados ni manipulados empíricamente, ni son producto de una relación entre otros atributos observables, la medida indirecta que persigue la psicometría no sería equiparable a la medida fundamental o derivada de la física y, por lo tanto, resultaría inviable. Esta fue la conclusión de la comisión organizada por la British Association for the Advancement of Science, que, como ya hemos comentado, lejos de cerrar la controversia se convirtió en uno de los debates más importantes para la medida indirecta de los fenómenos psicológicos mediante test.

En primer lugar, el modelo representacional fue desarrollado a partir de los trabajos de Stevens (1946) y Suppes (1951), y define la medida como el proceso de asignación de números a objetos a partir de unas reglas, de modo que reflejen relaciones empíricas entre los objetos. Estas relaciones, de manera comparable al modelo de Campbell, quedarían representadas por las propiedades de los números, pero no se centrarían exclusivamente en la aditividad. De hecho, Stevens abrió la puerta a otros tipos de relaciones para llevar a cabo la medida, aceptando por ejemplo la equivalencia o el orden. Medir no sería únicamente cuantificar, sino la representación numérica de los hechos empíricos en sentido amplio, y daría como resultado cuatro escalas clásicas para la psicología según el tipo de relación representada: escala nominal, ordinal, de intervalo y de razón.

En segundo lugar, el modelo operacional se basa en las contribuciones de Bridgman (1927) y Dingle (1950), a partir de las cuales se propuso la definición de cualquier concepto mediante las operaciones necesarias para medirlo. Así, de acuerdo con este modelo, la medida no sería otra cosa que el conjunto de operaciones necesarias para definir un concepto que, en última instancia, acaban produciendo números. Esta definición de la medida tiene puntos en común con la del modelo representacional, dado que asignar números –las puntuaciones obtenidas mediante tests– de acuerdo

con unas reglas es un tipo particular de las operaciones posibles, y los números resultantes no son otra cosa que el producto de las operaciones llevadas a cabo. En cambio, una importante diferencia entre estos dos modelos de medida se encuentra en el requisito según el cual los números representan o no un sistema de relaciones empíricas.

El modelo operacional, en cambio, limita el dominio de las operaciones a aquellas que de una manera consistente producen números y en ningún caso está interesado en la existencia de una realidad empírica que apoye estos números. Esta diferencia en el enfoque de la medida tiene además importantes implicaciones en cuanto al debate sobre la estadística permitida, dado que los partidarios del modelo operacional no ven la necesidad de establecer ninguna restricción a las operaciones estadísticas. No siendo el objetivo la búsqueda de relaciones empíricas comparables a las relaciones entre los números, sino los números en sí mismos, ninguna restricción puede ser aplicable a unas puntuaciones que, en palabras de Lord (1953), no tienen conciencia de dónde provienen.

Finalmente, el modelo clásico se propondría resolver esta discusión planteando la naturaleza cuantitativa de los fenómenos psicológicos como condición para la medida. Partiendo de los trabajos de Rozeboom (1966) y Jones (1971), y arraigado en el desarrollo de las teorías cuantitativas de la psicología de las décadas anteriores, este modelo define la medida como la determinación de la cantidad –del cuánto– en la que un atributo esté presente en el objeto medido. Es decir, medir consiste en determinar cuántas unidades están presentes en el atributo observado, pero a diferencia de los modelos de Campbell y Stevens, no exige la existencia de una relación empírica entre los objetos.

#### **Teoría clásica de los tests.**

La TCT es la teoría de los tests más extendida actualmente en la práctica de la psicometría y se basa en el modelo lineal clásico propuesto por Spearman (para una discusión histórica, podéis ver Traub, 1997), sistematizado por Gulliksen (1950) y reformulado posteriormente por Lord y Novick (1968), que articula el proceso de medida definiendo tres conceptos fundamentales: la puntuación verdadera, la puntuación empírica y el error de medida<sup>8</sup>. Para introducirlo, en este texto seguiremos la exposición de Muñiz (1996 y 2003).

Como cualquier tipo de medida científica, la medida indirecta de los fenómenos psicológicos se encuentra sujeta a estas variaciones o desviaciones no deseadas, y el modelo lineal clásico propone un conjunto de supuestos que permiten hacer una estimación de las puntuaciones verdaderas a partir de las puntuaciones empíricas obtenidas, y una definición de tests paralelos.

- Primer supuesto:  $V = E(X)$ .

La puntuación verdadera ( $V$ ) se define matemáticamente como la esperanza matemática de la puntuación empírica ( $X$ ). Es decir, el primer supuesto del modelo lineal clásico asume que, en caso de que fuera posible la administración de un test un número infinito de veces, la media de las puntuaciones empíricas obtenidas nos daría como resultado la puntuación verdadera del sujeto en el test. La puntuación empírica, por lo tanto, no es un sustituto de la puntuación verdadera, sino la mejor aproximación disponible en un proceso de medida que, de manera explícita, es reconocido como no libre de error.

- Segundo supuesto:  $U(X, e) = 0$ .

No existe correlación entre las puntuaciones verdaderas de los sujetos ( $X$ ) en un test y sus respectivos errores de medida ( $e$ ). Es decir, de acuerdo con el segundo supuesto no se espera que el tamaño de los errores cometidos esté sistemáticamente asociado al tamaño de las puntuaciones verdaderas.

- Tercer supuesto:  $U(e_j, e_k) = 0$ .

Si disponemos de dos tests diferentes ( $j$  y  $k$ ), no existe correlación entre los errores de medida cometidos con cada uno de ellos ( $e_j$  y  $e_k$ , respectivamente). Es decir, de acuerdo con el tercer supuesto, los errores de medida de los diferentes tests son aleatorios en cada ocasión y, por lo tanto, no se espera que exista ninguna relación entre ellos.

- Definición de tests paralelos:

$V_j = V_k$  y  $V^2(e_j) = V^2(e_k)$ . Finalmente, dos tests ( $j$  y  $k$ ) son paralelos siempre que sus puntuaciones verdaderas ( $V_j$  y  $V_k$ ) y sus varianzas de los errores de medida ( $V^2(e_j)$  y  $V^2(e_k)$ , respectivamente) sean idénticas.

### **El proceso de inferencia psicométrica.**

Antes de abordar sus especificidades, es importante tener presente que la medida objetiva y estandarizada de una muestra de comportamientos se ajusta, de manera general, al procedimiento establecido por el método científico. Lo podríamos resumir así:

- 1) Formular una pregunta de investigación o una hipótesis relevante.
- 2) Especificar y definir adecuadamente todas las variables involucradas.
- 3) Desarrollar o elegir los instrumentos y procedimientos necesarios para llevar a cabo las medidas.
- 4) Evaluar el funcionamiento de los instrumentos y procedimientos para obtener las garantías suficientes sobre la calidad del proceso de medida.
- 5) Recoger las evidencias necesarias que permitan responder a los objetivos de la investigación.

6) Resumir y, siempre que sea posible, tratar estadísticamente los datos obtenidos para determinar hasta qué punto los resultados son significativos y, por lo tanto, no son producto del azar.

La psicometría ofrece un conjunto de teorías, métodos y técnicas vinculadas al desarrollo y la administración de tests que dan el soporte necesario a los puntos 2, 3 y 4 del método científico cuando se utiliza en el contexto de la medida indirecta de los fenómenos psicológicos no observables. Y lo hace, como hemos señalado anteriormente, partiendo de una definición precisa de los fenómenos objeto de medida y de una selección de los comportamientos implicados que serán empleados como evidencias observables. Aun así, es importante hacer énfasis en algunas dificultades específicas que, como disciplina científica, la psicometría debe afrontar para garantizar la confianza en las inferencias establecidas sobre los fenómenos psicológicos no observables a partir de las puntuaciones obtenidas mediante tests. Todas ellas, en íntima relación con el reto esencial de la psicometría, son consecuencia directa de la dificultad añadida que supone la imposibilidad de observar y manipular directamente los fenómenos psicológicos objeto de interés. En este sentido, entre los retos específicos podemos señalar algunos de los más importantes:

- La medida indirecta de los fenómenos psicológicos no es unívoca. Como hemos ido viendo, no existe una única manera de definir el proceso de medida, así como no existe tampoco una única manera de entender los propios fenómenos objeto de medida. Son muchas las decisiones que los profesionales interesados en el desarrollo y la administración de tests deben tomar durante este proceso y, por lo tanto, son muchas las posibles soluciones a los problemas que surgen en la aproximación indirecta a la medida de los fenómenos psicológicos. Entre estas decisiones destacan las relacionadas con la elección de lo que es relevante medir, la selección y el muestreo de los comportamientos observables vinculados y la definición de las características que los tests deben tener para suscitarlos y medirlos adecuadamente.
- La teoría desempeña un papel fundamental en la medida psicométrica. Todas las decisiones que toman los profesionales interesados en el desarrollo y la administración de tests han de estar fundamentadas, como hemos señalado anteriormente, por dos tipos de teorías. Por un lado, por una teoría sustantiva muy establecida en torno a los fenómenos psicológicos y los comportamientos de las personas, que servirá de contexto de referencia para el proceso de medida. Por otro, una teoría de los tests que permitirá adoptar el enfoque más adecuado para llevar a cabo la medida y establecer la relación entre las puntuaciones obtenidas y los fenómenos psicológicos no observables, objeto de interés. Si los tests tienen sentido y sirven como instrumentos de medida en el campo de la psicología es, precisamente, gracias al apoyo que estos dos tipos de teorías proporcionan durante todo el proceso.
- La selección y el muestreo de los comportamientos impone importantes limitaciones. Un tercer reto importante para la medida indirecta que se propone

la psicometría es la selección y el muestreo de los comportamientos empleados como evidencias observables. En consonancia con la teoría sustantiva, el desarrollo y la administración de tests parten de una definición que orienta hacia unos comportamientos vinculados y no a otros, que, a la vez, deben ser adecuadamente muestreados dada la imposibilidad material de suscitar y medir el dominio entero de comportamientos al que pertenecen. De hecho, un comportamiento determinado puede ser empleado como evidencia de fenómenos psicológicos diferentes en función de la perspectiva teórica con la que se fundamenta el proceso de construcción de los tests. Importantes limitaciones vinculadas al tiempo requerido y las condiciones de administración también han de ser adecuadamente resueltas para hacer operativos los tests desde un punto de vista práctico.

- Las puntuaciones obtenidas requieren una interpretación adecuada. Las puntuaciones obtenidas mediante el uso de tests no son, per se, informativas y han de ser siempre interpretadas para responder al propósito con el que los tests han sido desarrollados. Ya sea desde un punto de vista general, con relación al contenido y su propósito, ya sea con relación al sistema de referencia empleado –según la norma, es decir, respecto a la ejecución de un grupo de referencia, o referidos a un criterio, esto es, respecto al grado de adecuación o consecución de este criterio–, el uso de tests requiere unos conocimientos y unas destrezas que solo podemos encontrar en manos de profesionales cualificados y que siempre debe cumplir con los objetivos más generales que persigue la evaluación psicológica a la que sirve.
- La medida psicológica no se puede entender si no es en relación con otras medidas o acontecimientos observables. Pese a la importancia de la teoría sustantiva en el desarrollo de los tests, no podemos obviar finalmente que la medida indirecta de los fenómenos psicológicos no tiene sentido solo en sí misma, sino que lo tiene en relación con otras medidas derivadas de la misma teoría u otros comportamientos observables que corroboren el proceso de medida mediante tests.

### **Construcción y administración de tests.**

Presentados los fundamentos de la psicometría, en este último apartado nos haremos cargo del proceso de construcción y administración de tests. Para hacerlo, empezaremos desarrollando las diferentes fases implicadas en el diseño y construcción de un nuevo test.

#### **El proceso de construcción de test.**

##### 1) Delimitación de la finalidad del test.

El desarrollo de un nuevo test empieza con una determinación clara del propósito para el que se pretende recoger información relevante en el contexto de la evaluación psicológica. Tal y como hemos señalado, los tests pueden servir a multitud de

finalidades, pero entre ellas podemos destacar algunas importantes, como son describir, clasificar, diagnosticar, explicar o hacer predicciones sobre el comportamiento de las personas.

## 2) Definición de los fenómenos psicológicos objeto de medida.

De acuerdo con el proceso de inferencia psicométrica, el segundo paso para la construcción de un test consiste en la delimitación precisa de los fenómenos que se pretenden medir. Para hacerlo, recordémoslo, es necesaria una teoría sustantiva sólida y bien contrastada sobre los fenómenos psicológicos y los comportamientos de las personas que servirá de referencia en el proceso de la medida mediante el nuevo test.

## 3) Selección y muestreo de los comportamientos observables.

Una vez establecido el fenómeno objeto de medida, la teoría sustantiva proporciona también el contexto necesario para elegir los comportamientos implicados que serán empleados como evidencias observables. Su representación adecuada es fundamental para no omitir ningún comportamiento relevante, así como para evitar incluir otros no directamente relacionados con el objeto de medida. En ocasiones resulta recomendable también llevar a cabo observaciones, entrevistas a informadores clave o grupos de discusión que proporcionen información complementaria sobre los comportamientos de interés a partir de las experiencias de los participantes.

## 4) Especificación de las características del test.

Delimitado el objetivo, definido el objeto de medida y seleccionadas las evidencias observables necesarias, el siguiente paso consiste en elegir las características del nuevo test para suscitar y medir adecuadamente los comportamientos de interés. En primer lugar, decidiremos si se trata de un test de habilidad o de potencia, en el que se evaluarán el acierto y el error en las respuestas para determinar la competencia, la aptitud o los conocimientos de los individuos, o de una prueba para evaluar motivaciones, preferencias, opiniones o actitudes. Es decir, si se trata de una prueba de ejecución máxima o de ejecución típica, como también son conocidas. De acuerdo con la clasificación de los tests que hemos hecho en esta introducción a la psicometría, es el momento también para decidir el formato del nuevo test, optando por una prueba de lápiz y papel, de manipulación o de medidas fisiológicas.

## 5) Desarrollo de los ítems que conformarán el test.

Una vez especificadas las características generales del nuevo test, en esta fase se llevará a cabo el desarrollo de los elementos que contendrá. Esta no es una tarea sencilla y generalmente supone la colaboración de un grupo de expertos en el campo para encontrar la mejor representación de las muestras de comportamiento seleccionadas.

#### 6) Elección de una teoría de los tests.

Cerrando el círculo establecido en el proceso de inferencia psicométrica, la siguiente fase en el proceso de construcción implica la elección de una teoría de los tests que, mediante un modelo de medida psicométrica determinado, servirá para relacionar los fenómenos psicológicos objeto de interés y las puntuaciones obtenidas mediante los ítems que conforman el nuevo test. Esta elección es capital teniendo en cuenta sus consecuencias en los métodos y técnicas empleadas posteriormente para evaluar las propiedades de los ítems y del test en su conjunto con el objetivo de garantizar la confianza en las inferencias establecidas sobre el comportamiento de las personas a partir de las puntuaciones obtenidas.

#### 7) Realización de una prueba piloto.

Una vez construido el test, incluyendo los ítems potenciales y el formato de las respuestas, es necesario redactar las instrucciones que lo acompañarán y definir las condiciones en las que será administrado. Una prueba piloto servirá para evaluar el grado de comprensión de estas instrucciones, la viabilidad de las condiciones para administrar el test, detectar posibles dificultades en el momento de registrar las respuestas y llevar a cabo un primer análisis de las propiedades, tanto de los ítems como del test en su conjunto, de acuerdo con los métodos y las técnicas indicadas por la teoría de los tests utilizada.

#### 8) Desarrollo del estudio de campo.

Una vez establecido el test definitivo, el siguiente paso consiste en su administración a la población a la que se dirige. Con esta información se desarrollan las normas o baremos para permitir la interpretación de las puntuaciones en relación con la ejecución del grupo de referencia. En el caso de que se trate de un test referido a criterio, en lugar de normas o baremos se determinan los puntos de corte que permitirán distinguir los diferentes grados de adecuación o consecución del criterio. Asimismo, se profundizará en el trabajo sobre las propiedades psicométricas de los ítems y del test en su conjunto, atendiendo especialmente a la fiabilidad y validez de las medidas obtenidas. Esta es una parte fundamental del proceso de construcción de test y es la que en última instancia garantiza que la medida psicométrica cumple con todas las exigencias científicas.

#### 9) Elaboración del manual del test.

Con toda esta información se desarrolla la documentación que acompañará al nuevo test, donde se ha de incluir información relevante relativa a las diferentes fases involucradas en su construcción: fundamentación teórica, finalidad y población a la que se dirige, instrucciones para la administración, información para la interpretación de las puntuaciones obtenidas y análisis de las propiedades psicométricas. Las primeras publicaciones en revistas científicas sirven para empezar a difundir toda esta

información, que como ya hemos comentado no siempre acaba constituyendo un manual propiamente dicho.

#### 10) Revisión y mejora del test.

Con la publicación del test, ya sea licenciándolo o difundiendo libremente, el test se pone a disposición de la comunidad científica para obtener nuevas evidencias, que, mediante el trabajo independiente de diferentes investigadores, servirán para mejorar el conocimiento sobre su funcionamiento y sus propiedades psicométricas, así como para adaptarlo a otros entornos socioculturales o a otras poblaciones diferentes para las que ha sido desarrollado el nuevo test. Cambios en los fenómenos psicológicos objeto de medida, nuevos avances en la aproximación teórica a estos fenómenos o la adaptación del test a nuevas condiciones de administración o poblaciones de interés son algunas de las razones que justifican un trabajo de revisión y actualización prácticamente indefinida para refinar su funcionamiento y valorar adecuadamente su utilidad al servicio de los objetivos de la evaluación psicológica.