



Nombre de alumno: José Emmanuel Gómez Mejía.

Nombre del profesor: Luis Miguel Sánchez Hernández

Nombre del trabajo: Resumen.

Materia: Psicometría.

Grado: 3°.

Grupo: Psicología.

Aproximación histórica y conceptos básicos de la psicometría

La psicometría se propone dar al lector algunas claves importantes para abordar la complejidad de los conocimientos, procedimientos y valores vinculados con el desarrollo y la administración de tests. Partiendo del encuadre de la psicometría en el contexto general de la psicología, empezaremos desarrollando una aproximación histórica a su nacimiento, desde los antecedentes remotos hasta los primeros desarrollos, describiremos el estatus actual de la psicometría como disciplina científica y, finalmente, construiremos una definición formal teniendo en cuenta el papel que desempeñan los tests como instrumentos de evaluación psicológica. A continuación, abordaremos los fundamentos de la psicometría, donde tendremos la oportunidad de ofrecer una definición y clasificación de los tests, abordar los modelos de medida psicométrica, introducir la teoría clásica de los tests y recapitular el proceso de inferencia psicométrica en el que se basan los tests. Finalmente, discutiremos las cuestiones relativas a la construcción y administración de tests, tratando las diferentes fases implicadas en el desarrollo y ofreciendo algunos criterios importantes para la valoración de los tests disponibles en la literatura. Concluiremos esta introducción con una discusión de los aspectos éticos y deontológicos vinculados al uso de tests en la práctica profesional de la psicología.

1. La psicometría en el contexto de la psicología

Entender un campo de estudio determinado implica, en la mayoría de las ocasiones, conocer sus raíces fundacionales. Es decir, conocer los problemas a lo que se enfrentaron los pioneros, la manera como los entendían y las soluciones que les dieron. A pesar de que este no es el lugar para hacer un viaje en profundidad¹, 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 tests 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.

1.1. Una aproximación histórica a la psicometría

La aproximación histórica a la psicometría se remonta al siglo XIX, cuando los psicólogos comenzaron a interesarse por la medición de los aspectos cuantitativos de la mente humana.

Francis Galton a finales del siglo XIX, Galton fue pionero en el desarrollo de técnicas de medición psicológica. Propuso el concepto de "herencia de la capacidad mental" y desarrolló pruebas de habilidades mentales para evaluar la inteligencia.

Alfred Binet a principios del siglo XX, Binet y su colaborador Simon desarrollaron la primera escala de inteligencia, conocida como la Escala Binet-Simon. Esta escala se utilizaba para identificar a los niños con retraso mental y fue la precursora de los modernos tests de inteligencia.

Charles Spearman en 1904, Spearman propuso la teoría del factor g, que postulaba que la inteligencia está compuesta por un factor general (g) y factores específicos. Esta teoría sentó las bases para el desarrollo de los tests de inteligencia basados en la teoría de los factores.

Louis Thurstone en la década de 1930, Thurstone introdujo el modelo de la teoría de los tests, que se basaba en la idea de que los tests medían múltiples habilidades mentales o "factores primarios". Su enfoque permitió una comprensión más detallada de las habilidades cognitivas subyacentes a la inteligencia.

Raymond Cattell en la década de 1940, Cattell propuso la teoría de los dos factores, que distinguía entre inteligencia fluida (capacidad para razonar y resolver problemas nuevos) e inteligencia cristalizada (conocimiento adquirido y habilidades prácticas). Esta distinción tuvo un gran impacto en la psicometría y en la comprensión de la inteligencia.

David Wechsler en la década de 1950, Wechsler desarrolló las Escalas Wechsler de Inteligencia para adultos y niños. Estas escalas se convirtieron en las pruebas de inteligencia más utilizadas y proporcionaron una medida más completa y comprensiva de la inteligencia, incorporando diferentes habilidades cognitivas.

Teoría de respuesta al ítem (TRI) a partir de la década de 1960, se desarrolló la TRI, que es un enfoque estadístico para modelar la relación entre las respuestas de los individuos a los ítems de un test y las habilidades o características que se están midiendo. La TRI permitió una mejora en la precisión de las pruebas y una comprensión más profunda de las habilidades evaluadas.

1.2. La psicometría hoy

La psicometría, un nuevo espacio de trabajo metodológico en torno al desarrollo y a la administración de tests, 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. Sin embargo, hasta la década de los años treinta del siglo XX no podemos situar los inicios de su constitución como disciplina científica² tal y como la conocemos en la actualidad. Es importante empezar señalando de nuevo el papel decisivo de Thurstone, quien, con el objetivo de establecer y promover la psicología como una ciencia cuantitativa (Samejima, 2000), en 1935 fundó y fue el primer presidente de la Psychometric Society. Además, fue el impulsor de la primera revista especializada que todavía es de referencia obligada, la revista Psychometrika, que publicó algunos de los trabajos más importantes sobre los que se formalizó la psicometría. Pocos años después, en 1946, fue también el primer presidente de la División de Evaluación y Medida de la American Psychological Association. En este sentido, podríamos definir 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.

La psicometría en la actualidad desempeña un papel fundamental en la medición y evaluación de variables psicológicas. Mediante técnicas estadísticas y teorías psicométricas, se desarrollan tests confiables y válidos que permiten una comprensión más precisa de los constructos psicológicos. La psicometría también se adapta a los avances tecnológicos y se enfrenta a desafíos éticos y culturales para garantizar la validez y la equidad en la evaluación psicológica.

1.3. La psicometría en el contexto de la evaluación psicológica

Una vez recorrido el camino de los antecedentes y los orígenes de la psicometría moderna, estamos en disposición de abordar formalmente su definición. Y podemos hacerlo a partir de las definiciones que se han ido proponiendo en los textos de referencia. Sin embargo, esta tarea no está exenta de riesgo. Los autores elaboran definiciones concretas que, probablemente, no reflejan todos los matices con los que después las contextualizan en sus textos. A pesar de esto, podemos aprovechar estas propuestas para valorar los diferentes aspectos en los que se pone un énfasis especial. Al fin y al cabo, no cabe todo en una definición. Más allá del tópico que dice que hay tantas definiciones como autores, podemos distinguir tres grandes aproximaciones. Por un lado, 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.

La psicometría es una rama de la psicología que, mediante teorías, métodos y técnicas vinculados al desarrollo y la administración de tests, se ocupa de la medida indirecta de los fenómenos psicológicos con el objetivo de hacer descripciones, clasificaciones, diagnósticos, explicaciones o predicciones que permitan orientar una acción o tomar

decisiones sobre el comportamiento de las personas en el ejercicio profesional de la psicología.

2. Fundamentos de la psicometría

Una vez abordada la psicometría desde una perspectiva histórica, tratando sus raíces y su construcción como disciplina científica en el contexto de la psicología, en este segundo apartado revisaremos sus fundamentos. Empezaremos discutiendo una definición y clasificación de los tests, progresaremos por los diferentes modelos de medida psicométrica desarrollados en las diferentes teorías de los tests y realizaremos una breve introducción a la teoría más extendida en la práctica actual: la teoría clásica de los tests. Por último, recapitularemos estos fundamentos atendiendo al proceso que la psicometría sigue para establecer sus inferencias sobre los fenómenos psicológicos no observables a partir de las puntuaciones obtenidas mediante tests. Empezamos, pues, con una definición y clasificación de los tests.

2.1. Definición y clasificación de los tests

los tests y los diferentes términos con los que se hace referencia a ellos nos lleva a fijarnos en dos grandes maneras de tratar las respuestas obtenidas para medir los fenómenos psicológicos de interés. Por un lado, los denominados tests de habilidad o de potencia tienen como objetivo evaluar la competencia, la aptitud o los conocimientos de los individuos a partir del acierto o la calidad de sus respuestas. Son pruebas que discriminan respuestas correctas e incorrectas y esta es la base para puntuar las ejecuciones individuales. Por otro lado, los denominados tests de personalidad tienen un objetivo diferente y pretenden conocer de manera general las motivaciones, preferencias, opiniones o actitudes de los individuos frente a un determinado estímulo. Este segundo tipo de tests no tienen respuestas correctas y por lo tanto no sirven para evaluar el acierto o el error de los individuos. Esta distinción es relevante desde un punto de vista terminológico, dado que los tests de personalidad son muchas veces llamados cuestionarios, inventarios o escalas, a pesar de que este uso no es siempre consistente. De hecho, el término escala puede resultar sinónimo de un test cuando está conformado por diferentes partes, o incluso de cada una de estas partes para reflejar las dimensiones o características específicas que miden en el contexto del test global. Por otro lado, otro término también extendido es el de batería, que tampoco tiene un significado unívoco y puede hacer referencia tanto a un test

compuesto por varias partes, como a una selección de diferentes tests administrados conjuntamente por un profesional en una evaluación psicológica determinada.

- 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. Otros propósitos más específicos son también posibles, como ya hemos comentado en la discusión de la definición formal de la psicometría.
- 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 computarizados, 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. También podemos distinguir los tests verbales y no verbales, en función de si se hace o no una presentación oral o escrita del test y sus instrucciones.
- Según el tratamiento de las respuestas. Siguiendo la diferencia que establecíamos entre los tests de habilidad o de potencia y los tests de personalidad, 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. Tal y como ya hemos comentado, 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. Finalmente, 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. En algunas ocasiones esta información queda reducida a la publicación de un artículo científico en el que el autor lo introduce, realiza una primera administración y presenta sus propiedades psicométricas.

2.2. Modelos de medida psicométrica

Todo tipo de medida científica, y la medida indirecta de los fenómenos psicológicos mediante tests no es una excepción, se fundamenta en una definición del propio concepto de medida. En este sentido, tal y como ya hemos introducido, en el núcleo de las teorías de los tests podemos encontrar un conjunto de modelos que establecen qué es exactamente medir y, como consecuencia, especifican la relación existente entre las puntuaciones empíricas obtenidas y los fenómenos psicológicos no observables objeto de interés. 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. Como veremos a continuación, suponen un importante punto de partida a partir del cual la psicometría ha desarrollado sus propios modelos de medida para abordar específicamente los fenómenos psicológicos no observables. De acuerdo con el trabajo de Campbell (1920), físico y miembro de la comisión que elaboró este informe, el modelo general de medida desarrollado en la física se basaría en la existencia de una equivalencia –un isomorfismo, en sus palabras– entre algunos atributos físicos y las propiedades aditivas de los números. Medir consistiría en asignar números a estos atributos para representar la aditividad de los objetos físicos y esta representación solo sería posible en cuanto que se satisficieran dos condiciones: que el atributo se pudiera ordenar, es decir, que fuera posible determinar que el atributo presente en un objeto es menor, igual o mayor que otro; y que se pudiera demostrar empíricamente la existencia de la aditividad mediante la manipulación empírica. Es lo que denominó medida fundamental, definiendo la medida de los atributos físicos por analogía con los números a partir de las operaciones llevadas a cabo sobre los objetos para cuantificar sus atributos.

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.

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.

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.

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. De hecho, el modelo clásico mide atributos, no objetos, y por lo tanto no contempla en su núcleo la demostración de la aditividad mediante la observación y la manipulación empírica de los objetos físicos. Es más, a diferencia de los modelos representacional y operacional, la

medida no se considera la asignación de números a objetos de acuerdo con unas reglas ni el resultado de un tipo particular de operaciones que produce números, respectivamente, sino que se define como el proceso de descubrimiento de relaciones numéricas entre los valores observados en un atributo cuantitativo.

2.3. 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. La TCT resulta una aproximación útil para la estimación de las puntuaciones verdaderas a partir de las puntuaciones empíricas, pero no está exenta de algunas limitaciones, que no son fáciles de resolver desde el modelo lineal clásico. Tal y como discutieron Lord y Novick (1968) en su reformulación, esta teoría tiene dos dependencias importantes en relación con los instrumentos y las muestras utilizadas. Por un lado, una dependencia de las puntuaciones empíricas obtenidas respecto a los instrumentos, de manera que, por ejemplo, dos tests de inteligencia independientes utilizan escalas diferentes y, por lo tanto, las puntuaciones que se obtienen no resultan comparables directamente. Por otro lado, una dependencia de las propiedades psicométricas de los tests respecto a las muestras utilizadas para desarrollarlos, de manera que, por ejemplo, la dificultad de los ítems que conforman un test depende de las características de los individuos a quienes se administra el test. Tratando de resolver estos problemas, la TRI se propone cambiar el foco desde el tratamiento del test entero al tratamiento individual de los ítems.

2.4. El proceso de inferencia psicométrica

Para concluir con sus fundamentos, recapitularemos el proceso que la psicometría sigue para establecer sus inferencias sobre los fenómenos psicológicos no observables a partir de las puntuaciones obtenidas mediante tests. 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.

- 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.

De acuerdo con este procedimiento general, 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.

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 de los fenómenos psicológicos no está libre de error. El error es un componente inherente a cualquier proceso de medida y debe ser objeto de un tratamiento adecuado para conocerlo y minimizarlo. La medida indirecta añade una dificultad específica que, de manera sintética, encontramos en la imposibilidad de observar directamente tanto los fenómenos psicológicos objeto de interés como el error cometido durante el proceso de medida. Partiendo de las puntuaciones obtenidas, y gracias al apoyo de los modelos de medida desarrollados en las diferentes teorías de los tests, la psicometría dedica una buena parte de sus esfuerzos a conseguir la precisión necesaria que, desde un punto de vista científico, es también exigible a la medida indirecta de los fenómenos psicológicos.

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. Como hemos dicho anteriormente, la medida mediante tests debe proporcionar información relevante con el objetivo de hacer descripciones, clasificaciones, diagnósticos, explicaciones o predicciones que permitan orientar una acción o la toma de una decisión sobre el comportamiento de las personas, y por esa razón las puntuaciones obtenidas mediante tests requieren un apoyo externo a su propia formulación para demostrar su utilidad práctica.

Estos principios sirven para evaluar las propiedades psicométricas de los ítems y de los tests en su conjunto, y entre estos principios básicos destacan los siguientes:

- Fiabilidad: la precisión con la que los tests llevan a cabo la medida.
- Validez: la confianza en que las medidas se corresponden realmente con lo que se proponen medir.

Los diferentes métodos y procedimientos vinculados a estos principios son la clave del éxito de los tests como instrumentos de evaluación psicológica y serán, por lo tanto, objeto de un tratamiento en detalle en capítulos posteriores.

3. 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. A continuación, y partiendo de las claves que nos proporciona el conocimiento de este proceso de construcción, discutiremos algunos criterios importantes para evaluar las características de los tests disponibles en la literatura y valorar su conveniencia en relación con los objetivos de la evaluación psicológica a la que deben servir. Finalmente, abordaremos los aspectos éticos y deontológicos vinculados al uso de tests en el contexto general de la práctica profesional de la psicología.

3.1. El proceso de construcción de tests

El desarrollo de instrumentos de medida es un proceso fundamental para cualquier disciplina científica. Y no lo es menos para la psicometría, en la que, como ya hemos discutido ampliamente, la imposibilidad de observar y manipular los fenómenos objeto de interés añade una importante complejidad.

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. Un propósito bien delimitado es la primera condición que hay que cumplir para garantizar el éxito en la construcción del nuevo test y permitirá su introducción en el proceso general de evaluación psicológica.

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. Este marco teórico es fundamental en el momento de planificar una representación adecuada de los fenómenos psicológicos y permite conocer en detalle otros instrumentos desarrollados previamente en la investigación psicológica. Evaluar sus limitaciones y sus puntos fuertes supone una buena guía para enfocar el trabajo necesario para dar soporte teórico al 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. Como ya hemos discutido, esta no es una fase menos importante en cuanto a que es la base para el correcto desarrollo de los ítems que conformarán el nuevo test.

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. Asimismo, se deben preparar los materiales necesarios y se determinará el tipo de administración más adecuada, eligiendo básicamente entre una prueba individual o de administración en grupo. Finalmente, también se decidirá el método más adecuado para interpretar las puntuaciones obtenidas, ya sea en base a la norma o en referencia a un criterio.

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. De hecho, no es poco habitual desarrollar muchos más ítems de los estrictamente necesarios para evaluar su comportamiento en el marco del test y seleccionar los más idóneos con relación al propósito del nuevo test. El objetivo final es disponer de tantos como sean necesarios para representar adecuadamente las diferentes dimensiones de los fenómenos psicológicos objeto de interés. Este es el momento también

para decidir el formato que adoptarán las respuestas, que servirán para codificar los comportamientos de manera estructurada y de acuerdo con unas reglas claras.

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. Es el momento para valorar su comportamiento en el proceso de medida y, a partir de esta información, refinar el test modificando, descartando o añadiendo nuevos ítems. Pruebas adicionales pueden ser requeridas para evaluar correctamente las modificaciones introducidas antes de cerrar esta fase y proceder con el desarrollo del estudio de campo final.

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. Para hacerlo, se selecciona la muestra de participantes necesarios, que, como ya hemos señalado, puede ser probabilística o no probabilística en función de si queremos representar la población de referencia o evaluar a un grupo de personas que cumplen unas determinadas características. 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 tests 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 difundiéndolo 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.

3.2. Criterios para la valoración de tests

Construir un test nuevo no es la práctica más habitual en el ejercicio profesional de la evaluación psicológica. De hecho, como hemos visto, es un trabajo laborioso que requiere un importante esfuerzo por parte de varios profesionales durante un periodo de tiempo prolongado para planificar, ejecutar y analizar los datos recogidos en uno o más estudios de campo. Por ello la mayoría de las situaciones en las que se utilizan los tests no comienzan necesariamente con la creación de uno original. En cambio, parten de la búsqueda y valoración de algún test existente en la literatura que encaje con los objetivos de la evaluación psicológica.

3.2. Aspectos éticos y deontológicos en el uso de tests

Como hemos discutido ampliamente a lo largo de esta introducción a la psicometría, el uso de tests como instrumentos de evaluación psicológica es una de las prácticas habituales en el ejercicio de la psicología. Sea al servicio de la intervención en los diferentes campos profesionales, sea en el centro del desarrollo de los campos académicos y científicos, los tests son unos instrumentos indispensables para la medida objetiva y estandarizada de los fenómenos psicológicos. Y como tales se encuentran sujetos a una reflexión sobre las consecuencias éticas y morales que implica su uso. Es decir, en cuanto que instrumentos, pueden ser utilizados o no de manera adecuada y de acuerdo con unas finalidades u otras. No es este el lugar para hacer una discusión filosófica en profundidad sobre lo que significa actuar bien o hacer las cosas correctamente, pero sí para concluir esta introducción reflexionando sobre algunas cuestiones importantes que tienen que ver con un uso de los tests responsable, justo y respetuoso con los derechos y la dignidad de las personas.

entre los principios generales que deben regir la actividad profesional de los psicólogos en el Estado español, plantea:

- 1) La atención primordial al bienestar, la salud, la calidad de vida y la plenitud del desarrollo de las personas.
- 2) La protección de los derechos humanos y la responsabilidad en la fundamentación objetiva y científica de las intervenciones profesionales.
- 3) El no desarrollo o contribución a prácticas que atenten contra la libertad y la integridad física y psíquica de las personas.
- 4) La obligatoriedad de informar al menos a los organismos colegiales ante el conocimiento de cualquier violación de los derechos humanos o tratamiento degradante.
- 5) El respeto por los criterios morales y religiosos de las personas.
- 6) La no discriminación por cualquier diferencia o motivo.
- 7) La denegación de cualquier beneficio o provecho que se pueda extraer como consecuencia de las relaciones de poder o superioridad establecidas.
- 8) El uso de un lenguaje prudente y crítico ante etiquetas despreciativas o discriminatorias en sus informes.

9) El respeto por la actividad de los otros profesionales y la libre competencia.

10) La denuncia de prácticas ilegítimas o intrusivas.

11) La imparcialidad ante los posibles conflictos de interés entre los psicólogos o las instituciones en las que desarrollan su actividad y los intereses de las personas.

Estas directrices establecen que los usuarios competentes de los tests deberían:

1) Responsabilizarse del uso ético de los tests:

- Actuando de manera profesional y ética
- Asegurando que tienen los conocimientos y las habilidades necesarias.
- Haciéndose responsables del uso de los tests.
- Manteniendo de forma segura los materiales de los tests.
- Garantizando la confidencialidad de los resultados.

2) Comprometerse con las buenas prácticas en el uso de los tests:

- Estudiando la utilidad de los tests en los procesos de evaluación
- Eligiendo tests bien fundamentados y apropiados para la situación.
- Atendiendo a las cuestiones relacionadas con la equidad.
- Preparando adecuadamente las condiciones de administración.
- Administrando correctamente los tests.
- Obteniendo las puntuaciones y analizándolas con exactitud.
- Interpretando los resultados correctamente.
- Comunicando los resultados de manera clara y exacta.
- Evaluando adecuadamente el funcionamiento y las propiedades de los tests.