



## Fundamentos Psicométricos en la Evaluación Psicológica

Laura Edna Aragón Borja<sup>1</sup>

Unidad de Evaluación Psicológica Iztacala  
Universidad Nacional Autónoma de México  
Facultad de Estudios Superiores Iztacala

### RESUMEN

La importancia de la evaluación psicológica es obvia: por una parte, es el primer paso que nos conduce a la planeación de la intervención y, por otra, nos sirve como punto de comparación para determinar el éxito de la terapia. El objetivo del presente artículo es exponer algunos conceptos clave que caracterizan a la evaluación psicológica y explicar cuáles son sus fundamentos psicométricos. Se discute así el concepto de medición a la luz de lo que se mide en psicología y cómo lo medimos y se hace una diferenciación entre medir conductas y medir atributos psicológicos. Se señala la importancia de la psicometría como una teoría de la medida en la creación de instrumentos de evaluación psicológica que garanticen su validez científica; se describen posteriormente dos de los requisitos indispensables que deben poseer los instrumentos de evaluación psicológica: confiabilidad y validez. Se explican las diferentes puntuaciones que arrojan las pruebas y cómo se interpretan, se señalan las partes que debe contener un reporte de evaluación psicológica y finalmente, se puntualizan algunas consideraciones éticas del proceso de evaluación psicológica.

Palabras clave: evaluación psicológica, medición psicológica, psicometría, confiabilidad, validez, puntuaciones z, decatipos, percentiles

---

<sup>1</sup> La correspondencia concerniente a este trabajo puede dirigirse a la Mtra. Laura Edna Aragón, Facultad de Estudios Superiores Iztacala. División de Investigación y Posgrado. Av. de los Barrios No. 1, Los Reyes Iztacala, Tlalnepantla, Edo. de México, 54090. Correo electrónico: [laragon@servidor.unam.mx](mailto:laragon@servidor.unam.mx)

---

## ABSTRACT

Psychological assessment is an important issue for at least two reasons: on one hand, it is the first step that leads us to intervention and, for other, it help us to determine the success of the therapy. The objective of the this paper is to show some theoretical concepts that characterize the psychological evaluation and to explain its psychometric foundations. It is discussed what the measurement is, what we measure in psychology and how we measure psychological attributes. It is pointed out the importance of the psychometrics as a theory in the creation of psychological instruments that guarantee their scientific validity; then, it is described two of the indispensable requirements that all the psychological instruments should possess: fiability and validity. The different punctuations the tests give us are explained and how they are interpreted, the differents elements that should contain a report of psychological evaluation are also pointed and finally, some ethical considerations of the process of psychological assessment are remarked.

Keywords: psychological assessment, psychological measurement, psychometrics, fiability, validity ; z, stein and centils punctuations.

La Evaluación Psicológica es considerada, en la actualidad, como aquella disciplina de la psicología científica que se ocupa de la exploración y análisis del comportamiento de un sujeto o grupo de sujetos humanos, a los niveles de complejidad que se estime oportunos (motor, fisiológico, cognitivo), con distintos objetivos básicos y aplicados (detección, descripción, diagnóstico, selección, orientación, predicción, explicación, intervención, cambio, valoración, investigación), a través de un proceso de toma de decisiones en el que se conjuntan, utilizan y aplican una serie de dispositivos, tests y técnicas de medida y/o evaluación (Fernández, 1996).

También se reconoce que la evaluación psicológica es un procedimiento multirrasgo, ya que se pueden evaluar diferentes atributos de la persona; multimétodo, puesto que se utiliza más de un método de obtención de información y/o multijueces, esto es, lo usual es que participe más de un evaluador a lo largo del proceso; se trata así de un concepto típicamente multidimensional (Garaigordobil, 1998).

Garaigordobil (1998) considera que, para que un proceso pueda considerarse como evaluación psicológica, debe reunir tres requisitos:

- a) La determinación de un contexto y un objetivo, esto es, un evaluador, un evaluado, una técnica o procedimiento de medida y una interpretación de los valores obtenidos
- b) La existencia de un marco teórico que permita la ubicación de las hipótesis
- c) Poseer un sistema de categorías que permita la clasificación del evaluado

La evaluación psicológica es un proceso indispensable en toda intervención, ya que por lógica, no podemos planear una intervención si antes no tenemos información acerca del problema del cliente, información que sólo nos la puede proporcionar una correcta evaluación.

El objetivo de este artículo es exponer algunos conceptos clave que caracterizan a la evaluación psicológica, explicar cuáles son los fundamentos psicométricos que subyacen a toda evaluación psicológica y proporcionar algunos

---

lineamientos generales tanto de la interpretación de las puntuaciones de las pruebas psicológicas como de la elaboración de un reporte de evaluación.

### ¿QUÉ ES LA EVALUACIÓN PSICOLÓGICA?

La evaluación en psicología, en un sentido estrecho, se refiere a diagnosticar, es decir, a detectar la presencia de un estado, un conocimiento, una patología, etc., de acuerdo al campo en donde se esté evaluando. En un sentido amplio, el término de evaluación no se restringe sólo a diagnóstico, sino a todo un proceso cuya meta consiste en identificar, rotular, seleccionar, plantear objetivos, intervenir, pronosticar y volver a evaluar si los objetivos planteados se alcanzaron; su finalidad estriba sobre todo en su utilidad para la toma de decisiones.

Desde este punto de vista entonces, la evaluación psicológica no se conforma únicamente con aplicar algún tipo de prueba o instrumento con la finalidad de conocer el estado actual de un sujeto o grupo de sujetos en cuanto a algún atributo o problema psicológico, sino que va más allá y éste sólo sería el primer paso de un proceso que involucra una toma de decisiones que comprende, a su vez, seleccionar y plantear objetivos de terapia, seleccionar las técnicas y estrategias para la intervención, evaluar continuamente para constatar que los objetivos planteados se van alcanzando así como para detectar posibles problemas en su consecución y evaluar para ver si las metas finales se han alcanzado. De este modo, la evaluación no representa por tanto un fin en sí misma sino que persigue un objetivo eminentemente práctico y como tal está al servicio de la disciplina: su objetivo fundamental no es la explicación sino la búsqueda de medidas prácticas de intervención.

La importancia de la evaluación psicológica entraña un concepto muy importante que se refiere a las inferencias que de ella pueden derivarse, inferencias basadas en las respuestas que se dan en la situación de evaluación. Como mínimo, el proceso de evaluación supone que la conducta a evaluar por medio de algún instrumento específico, es una muestra de los repertorios de conducta que el evaluado posee; representa lo que el individuo ha aprendido como resultado de su interacción con el entorno social.

---

Sin embargo, así como no todas las evaluaciones son iguales, tampoco lo son sus fines. Al respecto, podemos considerar que dentro del proceso de evaluación psicológica existen estrategias selectivas y estrategias modificativas; dentro de las primeras se busca la optimización a través de la selección adecuada ya sea de personas –en determinadas condiciones, elegir a las personas con un criterio de ejecución o clasificarla de acuerdo a dicho criterio- o de condiciones –se parte de unos individuos concretos y la estrategia consiste en buscar las condiciones idóneas que mejor se adapten a cada uno de ellos de acuerdo al criterio de optimización-. Por otra parte, en la estrategia modificativa se busca una forma de actuar buscando la optimización a través de la modificación tanto de las conductas de las personas como de las condiciones. Dentro de esta última estrategia se debe considerar que para que sea útil nos debe proporcionar los elementos suficientes para determinar tanto el método modificativo en cada caso concreto así como poder verificar la efectividad del mismo una vez utilizado (Pawlik, 1980).

A este tipo de diagnóstico que utiliza la estrategia modificativa –en oposición al que usa la estrategia selectiva y que presupone por tanto que el individuo posee la cualidad evaluada como un rasgo permanente de su conducta-, se le llama diagnóstico de proceso, en el que se considera que la conducta es susceptible de cambiar a través de diferentes procesos modificativos.

Sin embargo, sea el diagnóstico del tipo que sea, el objetivo final que persigue toda evaluación es la toma de decisiones. Esta toma de decisiones presupone la existencia de alternativas entre las cuales hay que elegir. Se podría considerar entonces a la intervención psicológica como todo aquello que al final de un proceso evaluativo se puede implementar como consecuencia de dicho proceso.

---

## MEDICIÓN

El objetivo final de toda ciencia es describir, explicar y predecir los fenómenos que estudia; para lograr este objetivo, la ciencia debe recoger y comparar datos que permitan fundamentar sus teorías. La medida es el componente esencial de este proceso; la medición es lo que le ha permitido a la psicología alcanzar el status de ciencia y no sólo ser considerada como una doctrina psicológica.

La medición es por tanto una actividad prioritaria tanto en la ciencia psicológica en general como en la evaluación psicológica en particular; el proceso de medición es el que le permite a una disciplina alcanzar el estatus de ciencia ya que, una vez que se pueden medir sus datos, se pueden comparar, replicar, establecer modelos matemáticos para su explicación y realizar investigaciones controladas. La pregunta a resolver sería: ¿qué medimos en psicología y cómo es que lo medimos?

### ¿QUÉ MEDIMOS EN PSICOLOGÍA?

La respuesta más sencilla a esta pregunta es: lo que medimos en psicología es la conducta, el comportamiento. Pero la respuesta no nos dejaría satisfechos ya que por ejemplo, no tiene el mismo nivel de dificultad el medir una conducta como “fumar”, “hacer berrinche”, que las conductas de “palpitaciones por nerviosismo” o “estudiar” o “estar triste”; en realidad, no podemos medir esas conductas de manera directa ni absoluta, y entonces, más que conductas, lo que medimos son “atributos de la conducta” o “atributos psicológicos”. Por otra parte, aunque conductas como “fumar” sea fácilmente medible, a la psicología le interesaría primordialmente el componente psicológico de la conducta de fumar ya que esta conducta tiene otros componentes: biológico, fisiológico, motor, etc.; así, de la conducta de fumar, nos interesaría el componente psicológico o atributo psicológico, por ejemplo, la dependencia psicológica hacia el cigarrillo, las cognitivas que el sujeto presenta tanto cuando se va a fumar el cigarro como cuando no puede hacerlo, otras conductas que desencadenan tal conducta, las consecuencias tanto agradables como aversivas del hecho de fumar, etc.

Para que quede un poco más clara la distinción entre medir un objeto o un atributo, digamos que cuando medimos una mesa (que es un objeto físico), más que la mesa como tal, medimos ciertos atributos de ella: su largo, su ancho, su altura; del mismo modo, cuando medimos personas o comportamientos, no los medimos totalmente, sino sólo ciertos atributos psicológicos que nos interesan: motivación, ansiedad, depresión, inteligencia, afrontamiento, personalidad, fumar, etc.

Sin embargo, a diferencia de la mesa en que los atributos largo, ancho o altura los podemos medir directamente por medio de un instrumento que mida longitud, muchos de los atributos psicológicos no los podemos medir directamente, sino sólo los podemos medir indirectamente, por eso se habla de que los atributos psicológicos se miden a través de indicadores; esto es, instrumentos que nos van a mostrar indirectamente el atributo que pretendemos medir.

## PSICOMETRÍA Y TEORÍA DE LA MEDIDA

Así como en las ciencias físicas la elaboración de instrumentos se deriva de la teoría en la que están inmersos y de una tecnología que permite medir sus objetos de estudio, los instrumentos de evaluación psicológica responden a una teoría de la medida; como dice Martínez (1996), los instrumentos de medición psicológica necesitan demostrar su utilidad y validez científica.

En psicología, la psicometría es la que garantiza que los instrumentos de medición elaborados cuenten con las garantías científicas para su uso. La psicometría es el conjunto de métodos, técnicas y teorías implicadas en la medición de variables psicológicas; estudia las propiedades métricas exigibles en las mediciones psicológicas y establece las bases para que dichas mediciones se realicen de forma adecuada. El objetivo de la psicometría es proporcionar modelos para transformar los hechos en datos con la finalidad de asignar valores numéricos a los sujetos, sobre la base de sus respuestas (Muñiz, 1992)

## TEORÍA CLÁSICA DE LOS TESTS

La psicometría es una rama relativamente muy joven de la psicología; se origina con los primeros trabajos de Spearman de 1904 a 1913 (Muñiz, 1992), cuyo objetivo era encontrar un modelo estadístico que fundamentase las puntuaciones de los tests y permitiera la estimación de los errores de medición asociados a todo proceso de medición. A este modelo se le conoce como el modelo lineal clásico de Spearman, y es el que todavía en la actualidad perdura en la Teoría Clásica de los Tests, y se expresa como:

$$X = V + e$$

en donde **X** es la puntuación empírica de un sujeto, **V** es la puntuación verdadera y **e** es el término de error. De aquí se sigue que la puntuación de un sujeto en una prueba es igual a la puntuación verdadera, que nunca se puede conocer, más el error de medición; que mientras más confiable sea la prueba y por tanto menor sea el error, la puntuación empírica se acercará más a la puntuación verdadera. Reconocer que toda medida psicológica, al igual que las medidas de la ciencia contienen un término de error, fue un gran avance para la evaluación psicológica, ya que entonces se empieza a considerar que no basta el juicio humano para determinar los atributos psicológicos, y que los instrumentos de evaluación psicológica deben intentar, al igual que los instrumentos de las ciencias físicas, medir con precisión.

Se reconoce así la necesidad primero de establecer y luego de medir la confiabilidad y la validez de los instrumentos psicológicos, así como las recomendaciones técnicas tanto para la elaboración de tests como para su uso, temas ignorados hasta ese momento. Entonces, para poder considerarlos adecuados y científicos, los instrumentos de todas las ciencias deben contar con dos requisitos indispensables y la psicología no es la excepción: la confiabilidad y la validez.



## CONFIABILIDAD

La confiabilidad en sentido amplio, es la exactitud, la precisión con que un instrumento de medición mide el objeto, en términos estrictos, la confiabilidad sería la ausencia de errores de medición utilizando un instrumento determinado; sin embargo, en psicología, es difícil utilizar dichos términos en la medición de atributos psicológicos. Como alternativa se utilizan como sinónimo de confiabilidad el de estabilidad de la medida y el de consistencia interna (Muñiz, 1992).

La estabilidad de la medida se refiere a que, al medir un atributo psicológico con un determinado instrumento, éste será confiable si, al evaluar a los mismos sujetos con el mismo instrumento, las medidas obtenidas en la segunda aplicación sean muy similares a las obtenidas en la primera aplicación, esto es, sean estables a través del tiempo, lo cual indicaría que los errores de medición serían mínimos y por tanto la confiabilidad sería aceptable, atribuyéndose las diferencias encontradas entre una medición y otra a los errores aleatorios asociados al proceso de medición y no al instrumento (Muñiz, 1992).

Por otro lado, la consistencia interna se refiere a que, en un instrumento dado, los reactivos que lo componen son consistentes entre sí en la forma en que están evaluando el atributo psicológico propuesto, esto es, en aquellos sujetos que puntúan alto en el atributo, los sujetos de manera individual puntuarán alto en aquellos reactivos que tienden a medir dicho atributo y puntuarán bajo en aquellos que no lo miden, siendo así consistentes los reactivos entre sí en la evaluación del atributo a evaluar.

## VALIDEZ

La validez es el grado en que un instrumento de medición mide realmente lo que se propone; es el grado de adecuación, significación y utilidad de las inferencias específicas que pueden derivarse a partir de las puntuaciones de los tests, teniendo

---

en cuenta que lo que se valida no es el instrumento, sino la interpretación de los datos obtenidos por medio de él (Martínez, 1996). En este sentido, la validez se refiere principalmente a la utilidad de los datos proporcionados por el instrumento, a las inferencias que podemos deducir de la ejecución del sujeto en una prueba.

Existen tres tipos de validez: de constructo, de criterio y de contenido, y de acuerdo a la validez que posea un instrumento determinado, en esa medida podemos realizar inferencias sobre la conducta posterior del sujeto evaluado (Aragón y Silva, 2002). Así, en la validez de constructo, que valida un rasgo teórico o constructo hipotético, se validan las cualidades o rasgos psicológicos que mide el test y se valida no sólo el rasgo sino la teoría sobre la que éste descansa; lo que implica que el instrumento mide un constructo o rasgo teórico derivado de toda una teoría psicológica, subyacente al instrumento de medición y a la medida misma, por ejemplo, inteligencia, personalidad.

Otro tipo de validez es la de criterio, que puede ser concurrente o predictiva, y se refiere a que la medida obtenida por un sujeto en la prueba nos va a ayudar a predecir, en el tiempo presente o a futuro, un comportamiento, es la eficacia de un test en la predicción de situaciones específicas; se compara la actuación del sujeto en un test con un criterio, con una medida directa e independiente de lo que está destinado a medir; debe existir evidencia de que hay una relación entre las puntuaciones del test y las del criterio, por ejemplo aptitudes para el aprendizaje, preferencias vocacionales, selección de personal, un diagnóstico clínico.

Finalmente la validez de contenido es el grado en que un conjunto de reactivos representan adecuadamente un dominio de conductas; es una muestra representativa de un contenido teórico de conocimientos, habilidades o conductas; mide una muestra representativa de un contenido teórico de conocimientos o habilidades, y en este tipo de validez se garantiza que los resultados del sujeto en la prueba, se puedan generalizar al universo de contenido que la prueba representa, por ejemplo, rendimiento escolar, hábitos de estudio, errores de tipo disléxico, ansiedad, habilidades sociales.

De esta manera, tanto la confiabilidad como la validez son dos requisitos indispensables que debe poseer todo instrumento de evaluación psicológica, lo cual nos garantizará el poder usarlo confiablemente y asegurar que las inferencias y decisiones que de ellos se deriven, serán las más apropiadas.

## CLASIFICACIÓN DE LOS TESTS

No todos los instrumentos de evaluación psicológica son iguales; podemos clasificarlos de acuerdo a su interpretación, a su forma de comparación, a su finalidad, a su modo de aplicación, a su forma de aplicación y al área del comportamiento que miden.

De acuerdo a la interpretación, las pruebas se dividen en pruebas objetivas y pruebas subjetivas. Las pruebas objetivas nos proporcionan tablas y medidas de comparación de los sujetos contra los cuales comparar su ejecución; dichas tablas y medidas nos permiten interpretar los resultados de una manera clara, objetiva, no ambigua y la interpretación no depende del juicio del evaluador. Así, la ejecución de un sujeto arrojaría los mismos resultados y por tanto sería interpretada de la misma manera por dos evaluadores diferentes, esto es, independientemente del evaluador (por ejemplo, prueba de inteligencia WISC-RM, prueba de personalidad 16PF, prueba MMPI, prueba de percepción visoespacial de FROSTIG).

Por otro lado, las pruebas subjetivas son aquéllas en que la interpretación de la ejecución del evaluado descansa en los juicios del evaluador, que a su vez dependen de la teoría sobre la que la prueba descansa; aunque hay líneas generales de interpretación, los resultados de un mismo sujeto podrían ser interpretados de diferente manera por diferentes evaluadores (por ejemplo, prueba de las Manchas de Tinta de Roscharch, Prueba del Dibujo de la Figura Humana; Test de la Casa, el Árbol y la Persona).

Por la forma de comparación, las pruebas pueden ser referidas a la norma o referidas al criterio. En las pruebas referidas a la norma, la ejecución de un sujeto se compara contra la ejecución de un grupo o de una población, esto es, contra un

grupo normativo. Por ejemplo, si al aplicar una prueba de inteligencia a un niño de 7 años éste obtiene un CI de 85, un percentil de 25 y una edad escalar de 6:4; esto significa que, comparado con la población de niños de 7 años, su CI se encuentra por debajo de la norma esperada para su edad -que debería estar entre 89 y 110-, que su ejecución en esta prueba se encontró por abajo del 75% de los niños de su edad y que su ejecución en este atributo es comparable a la de un niño de 6 años 4 meses, esto es, 8 meses por debajo de su edad cronológica.

En las pruebas referidas al criterio, cada individuo es comparado contra un criterio no normativo, contra un criterio arbitrario determinado de antemano, y la ejecución del sujeto se compara para observar qué tanto se aleja o se acerca al criterio establecido, criterio que si bien es arbitrario en el sentido de que no es normativo, se basa en la experiencia y en cuestiones teóricas. Por ejemplo, un profesor puede determinar que para acreditar su materia, los alumnos deben poseer un mínimo de conocimientos que equivalen a un siete de calificación; al trabajar con un niño con síndrome de Down, se puede establecer como criterio para pasar a un siguiente objetivo de lenguaje, que debe obtener el 75% de respuestas correctas en una prueba de lenguaje, etc.; así, al evaluar a ambos casos, se determinaría si alcanzan el criterio señalado y si no, qué tan lejos se encuentran de él para seguir trabajando en la consecución del objetivo.

Por su finalidad, las pruebas pueden ser utilizadas para detección, investigación, diagnóstico, selección, predicción, planeación de un tratamiento, intervención, valoración del cambio, etc., teniendo siempre presente que, independientemente del objetivo de su aplicación, la finalidad última de toda aplicación de pruebas es la toma de decisiones. Por su modo de aplicación, las pruebas pueden ser de lápiz y papel, orales, de manipulación o ejecución, mixtas y computarizadas. Por su forma de aplicación pueden ser individuales, colectivas y mixtas.

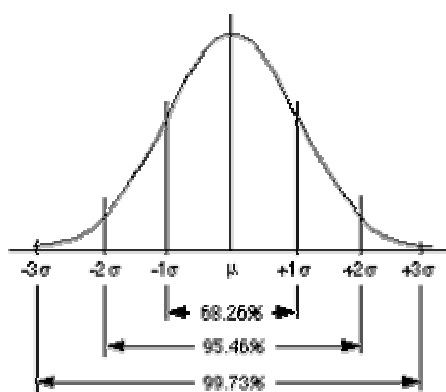
Finalmente, por los atributos que evalúan, pueden ser clasificadas, arbitrariamente, en pruebas del área educativa-intelectual (inteligencia, rendimiento, aptitudes, percepción, desarrollo psicológico, aptitudes para el aprendizaje, memoria, hábitos de estudio, lenguaje, etc.), pruebas del área clínica (personalidad, depresión,

motivación, ansiedad, habilidades sociales, adaptación, trastorno por déficit de atención, autoconcepto, socialización, autocontrol, afrontamiento, asertividad, etc.) y del área de selección y orientación vocacional (aptitudes vocacionales, preferencias e intereses profesionales y vocacionales, valores, etc.).

## NORMAS DE PUNTUACIÓN

Una vez que hemos elegido una prueba o seleccionado una batería de pruebas, de acuerdo al objetivo que persigamos con su aplicación, que desde luego estará muy ligado al motivo de consulta y al tipo de validez que proporcione la prueba, debemos tener en cuenta el tipo de puntuaciones que proporciona dicha prueba, y que son las que van a regir la interpretación. Las pruebas psicológicas nos pueden proporcionar puntuaciones típicas o  $z$ , decatipos, percentiles y, en algunos casos, edades escalares.

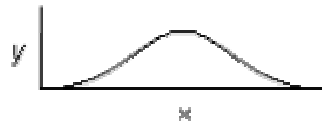
Las puntuaciones típicas o  $z$  son aquellas derivadas de la curva normal; esto es, si al aplicar una prueba a la población elegida, la distribución de las puntuaciones obtenida por los sujetos evaluados sigue una distribución normal, se aprovechan las propiedades de la curva normal para determinar las puntuaciones. Como el lector recordará, en una distribución normal, la media vale cero y la desviación estándar 1; así, mediante una transformación lineal, se iguala la media con una puntuación de 100 y la desviación estándar queda así de 15 o 16. Este tipo de puntuaciones se encuentran en las pruebas que miden inteligencia, desarrollo psicológico, habilidades psicolingüísticas, percepción.



Entre la media y una desviación estándar a la derecha y a la izquierda de la media, se encuentra el 68.26% de los casos (esto es, 34.13% de cada lado), entre la media y dos desviaciones estándar se encuentra el 95.44% de la población evaluada y entre la media y tres desviaciones estándar se encuentra el 99.72% de la población, encontrándose el 0.18% de la población por arriba de tres desviaciones estándares. De esta manera, cuando al aplicar una prueba psicológica en la población elegida, ésta sigue una distribución normal, las puntuaciones normativas para comparar posteriormente a cada individuo evaluado con su grupo de comparación, son la media más y menos una desviación estándar, lo cual significa que se le compara con la ejecución que realiza el 68.26% de la población para determinar lo adecuado o no de su comportamiento. Claro está que por ejemplo, si el individuo se encuentra por arriba de ese 68.26% esto no sería un problema sino por el contrario, una ventaja; por otro lado, aquel individuo que se encontrara por debajo de esa puntuación que obtiene el 68.26% de los casos, en realidad se encontraría por abajo del 84.08% de la población (abajo del 68.26% más los que se encuentran por arriba de ellos a la derecha de la curva normal), lo cual es un fuerte indicador de que existe un problema. Por ejemplo, si un niño de siete años no resuelve el número de problemas de aritmética que resuelven el 84-08% de niños de su edad, esto nos da pie para pensar que quizás a eso se debe su bajo rendimiento académico.

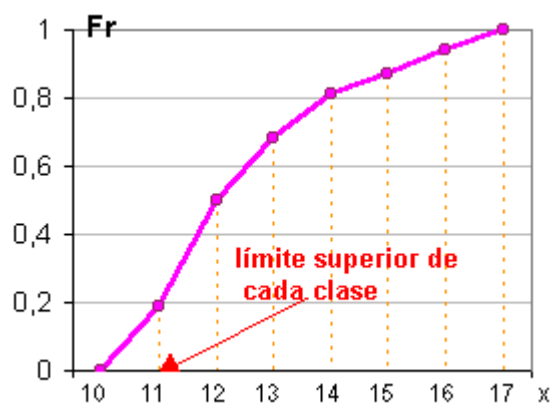
Los decatipos son también puntuaciones derivadas de una distribución de curva normal, sólo que, en lugar de tener la forma de campana de Gauss, es una curva normal planocúrtica, esto es, aplanada, en las que, en lugar de dividirse la curva en seis partes (la media más y menos tres desviaciones estándares), se divide en 10 para los decatipos (la media más y menos cinco desviaciones estándares), variando los porcentajes de población que se encuentran dentro de cada porción. En los decatipos 1 y 10 se encuentra el 2.3% de la población, en los decatipos 2 y 9 se encuentra el 4.4%, en los decatipos 3 y 8 el 9.2%, en los 4 y 7 el 15% y en los decatipos centrales, 5 y 6, se encuentra el 19.1% de la población. La interpretación se basa tomando como norma de comparación los cuatro decatipos centrales: el 4, 5, 6 y 7 y considerando entonces puntuaciones extremas los decatipos 1, 2 y 3 del lado izquierdo de la curva (que representarían al 15.9% de la población con puntuaciones

más bajas) y los decatipos 8, 9 y 10 a la derecha (el 15.9% con las puntuaciones más altas).



Un ejemplo de prueba que se distribuye de esta manera es la serie de pruebas de personalidad elaborada por R.B.Cattell: el cuestionario de 16 factores de personalidad o 16PF, el HSPQ, el CPQ y el ESPQ, que evalúan personalidad a personas mayores de 16 años, a adolescentes de 12 a 16 años, a niños de 8 a 12 años y a niños de 6 a 8 años, respectivamente. Al seguir las puntuaciones de esta prueba una distribución de curva normal de decatipos, Cattell interpreta los decatipos 1, 2 y 3, como el polo negativo del factor o rasgo de personalidad que está evaluando, por ejemplo: reservado, baja capacidad mental escolar, poca estabilidad emocional, sumisión, poco autocontrol, relajado; contrastándolos con el polo positivo que correspondería a los decatipos 8, 9 y 10: afectuoso, alta capacidad mental escolar, emocionalmente estable y maduro, dominante, bastante autocontrol, ansiedad; situando a las puntuaciones intermedias: 4, 5, 6 y 7 como poseedor del rasgo en términos medios, como lo poseen el 68.2% de la población, y que lo demuestran en mayor o menor grado según el contexto en el que se encuentren.

Sin embargo, como no es factible que toda prueba construida se distribuya normalmente al aplicarla a la población meta -que son la mayoría-, las puntuaciones que se utilizan son los percentiles (o centiles), que son las puntuaciones derivadas de una distribución de frecuencia acumulada y dividida en 100 partes, encontrándose en cada parte o percentil de la distribución el mismo número de casos, a diferencia de la curva normal, en que cada parte de la curva tiene un porcentaje diferente de población. Algunos percentiles importantes son la mediana y los cuarteles primero y tercero.



La mediana (el percentil 50) es el punto de una distribución de frecuencias acumulada que divide a la población en dos partes iguales, esto es, por arriba y por debajo de la mediana se encuentra el 50% de los casos. Los cuartiles dividen a la distribución de frecuencia acumulada en cuatro partes iguales, encontrándose en cada cuartil el 25% de la población; así, por debajo del cuartil uno (percentil 25), se encuentra el 25% de la población que obtuvo las puntuaciones más bajas y por arriba del tercer cuartil (percentil 75), se encuentra el 25% de la población que obtuvo las puntuaciones más altas.

De esta manera, si al aplicar una prueba a la población meta sus puntuaciones no se distribuyen normalmente, entonces es conveniente elaborar una distribución de frecuencias acumuladas y determinar los diferentes percentiles, tomando en cuenta que los percentiles de comparación serán el 25 y el 75, ya que entre ellos se encuentra el 50% de la población con puntuaciones intermedias, es decir, es la mitad de la población que se encuentra en medio de la distribución, y con ella se va a comparar la ejecución de los demás evaluados. Así, si un sujeto al aplicarle una prueba de inteligencia se encuentra en el percentil 60, esto querría decir que se encuentra por arriba del 60% de la población en el rasgo evaluado y que a su vez el 40% tiene una ejecución mejor que él; si se encuentra en el percentil 85, su ejecución se encuentra por arriba del 85% de los chicos de su edad y sólo por abajo del 15% de la población, y si su percentil es de 20, entonces esto indicaría que su ejecución se encuentra por abajo del 80% de la población normativa. Si estamos evaluando ansiedad, entonces encontrarse en el percentil 25 o menos no sería un problema, pero sí encontrarse en el percentil 75 o más arriba, ya que indicaría que el



sujeto evaluado sobrepasa al 75% o más de la población en ansiedad. Las normas aceptadas universalmente para determinar lo adecuado o no de un rasgo medido son el percentil 25 y 75, esto es, si el evaluado obtiene una puntuación percentil entre 25 y 75 se encuentra en la norma, acercándose a la norma alta de 60 a 75 y a la norma baja de 35 a 25; puntuaciones mayores a 75 serían consideradas altas e indicarían en qué porcentaje sobrepasa el evaluado a la población en el rasgo evaluado (inteligencia, ansiedad, depresión) y puntuaciones menores a 25 son consideradas bajas e indicarían al restarse de 100, el porcentaje por abajo del cual se encuentra el evaluado en el rasgo con respecto a la población de comparación (inteligencia, ansiedad, depresión).

Por último, las edades escalares, mentales o de desarrollo que establecen algunas pruebas, sobre todo las infantiles y algunas para adolescentes, proporcionan la edad promedio de ejecución del evaluado en el rasgo medido. Aquí se debe tener cuidado con la interpretación y no declarar que el evaluado tiene una edad mental de X años, sino especificar que sólo es en el rasgo evaluado. Por ejemplo, si al evaluar a un niño de 7 años con una prueba de inteligencia su CI fue de 84, su percentil de 20 y su edad escalar de 5:9, esto indicaría que se encuentra por debajo de la norma en cuanto a inteligencia, que su ejecución lo coloca por abajo del 80% de los niños de 7 años y que su ejecución en la prueba corresponde, en promedio, a la ejecución que tendría un niño de 5 años 9 meses, esto es, tendría un atraso aproximado en inteligencia de un año tres meses; pero de ninguna manera indicaría que el niño de 7 años se comporta como un niño de 5 años 9 meses en todas las demás esferas de su vida, sino que sólo fue su ejecución en la prueba la que se compara con esa edad.

De esta manera, si tenemos un cliente o un grupo de sujetos a quienes vamos a evaluar, en primer lugar, después de la entrevista y de conocer el motivo de consulta, determinamos claramente un objetivo de evaluación, seleccionamos la o las pruebas apropiadas para lograr ese objetivo, las aplicamos y calificamos de acuerdo con el manual de la prueba. De las puntuaciones directas o crudas obtenidas al calificar la prueba encontramos, en las tablas de los manuales de las pruebas utilizadas, las puntuaciones derivadas que pueden ser típicas, decatipos, percentiles y/o edades escalares, con lo cual podemos hacer la interpretación de lo

que significan esas puntuaciones, lo que representan para el sujeto en los atributos evaluados y a la luz de la teoría en la que está inmersa la prueba y el tipo de validez que tiene. El último paso pero no menos importante es la entrega de un reporte de evaluación, en el que plasmemos toda la información que hemos recabado y que sea de utilidad tanto a nosotros o al psicólogo que va a planear la intervención como al usuario para que pueda tomar decisiones.

## REPORTE DE EVALUACIÓN PSICOLÓGICA

Un reporte de evaluación debe tener los siguientes elementos:

**Carátula.** Con los datos personales del evaluado: nombre, dirección, teléfono, edad, grado de estudios y escuela a la que asiste.

**Motivo de la evaluación.** Describir muy claramente cuál es el motivo por el cual el usuario acude a solicitar una evaluación psicológica, quién lo remite y por qué y la información que se espera obtener al final del proceso.

**Antecedentes.** En este punto, se describe un resumen de la historia clínica y/o entrevista con el cliente y con otros significativos (padres, maestros, etc.); la información vertida en este punto debe ser pertinente y relevante con el motivo de la evaluación y, de acuerdo a quien(es) va(n) a leer el reporte, confidencial.

**Batería de pruebas elegidas.** De acuerdo al motivo de evaluación, el psicólogo elegirá las pruebas que considere pertinentes. En este punto se recomienda poner el nombre de la prueba con su respectivo autor, el objetivo que la prueba tiene (esta información viene en el manual de cada prueba), una breve descripción de la misma y sobre todo y muy importante, la justificación de su aplicación, esto es, por qué el psicólogo la ha elegido, qué información espera obtener de ella.

Por ejemplo, si a un niño con bajo rendimiento académico se decide aplicarle la prueba de inteligencia WISC-RM de Wechsler, cuyo objetivo es obtener el CI de niños en edad escolar, la descripción sería que esta prueba mide inteligencia a través de dos escalas: una verbal y una de ejecución; que la escala verbal consta de las subpruebas: información, semejanzas, aritmética, vocabulario, comprensión, y retención de dígitos, y que la escala de ejecución consta de las subpruebas: figuras incompletas, ordenación de dibujos, diseño con cubos, composición de objetos, claves y laberintos. La justificación podría ser que, con su aplicación, se estimaría si las habilidades intelectuales del niño son las esperadas para su edad y ver cómo están influyendo éstas en su desempeño escolar.

**Descripción, análisis e interpretación de resultados.** Una vez aplicadas y calificadas las pruebas elegidas, se puede describir la ejecución del niño en cada una, proporcionando los valores obtenidos, analizando esa información con respecto a las normas de la población de comparación e interpretando los resultados relacionándolos con el motivo de la consulta, con lo observado en las sesiones de evaluación y con la información previa y las entrevistas.

**Actitud durante la evaluación.** Aquí se describen aquellos aspectos relevantes relacionados con el proceso de evaluación y el motivo de la consulta que tuvieron lugar desde el primer momento en que se contactó con el evaluado hasta la terminación de la evaluación.

**Conclusiones.** Se trata de integrar toda la información obtenida durante el proceso, relacionando los diferentes resultados obtenidos en las pruebas aplicadas y el motivo de la evaluación y la entrevista, para ayudar en la toma de decisiones.

**Sugerencias.** Este último punto tiene que ver con la toma de decisiones; basadas en la información proporcionada en el reporte, se proporcionan sugerencias de intervención o tratamiento, tanto a largo plazo como a lo inmediato, que pudieran ayudar al sujeto en la resolución del problema encontrado.

**Firmas.** Firmar el reporte con el nombre completo, la firma y la adscripción de la institución, con el objetivo de darle al proceso la legalidad requerida, ya que puede

---

ser probable que ese reporte lo lea otro colega psicólogo o bien otro profesionista (médico, maestro, trabajadora social, juez, etc.).

A lo largo de este artículo hemos revisado qué es la evaluación psicológica, cuáles son sus fundamentos psicométricos, cómo se interpretan las puntuaciones que nos proporcionan las pruebas psicológicas y cómo elaborar un reporte de evaluación. Para finalizar, acotemos algunas consideraciones éticas del proceso de evaluación.

## ÉTICA

Algunos artículos importantes que reglamentan la actividad del psicólogo en lo referente a evaluación, están plasmados en el “Código Ético del Psicólogo”, que publicó la Sociedad Mexicana de Psicología (2000), y son el 3, 18, 24, 53, 58, 59 y 60. Se recomienda que el lector lea no sólo estos artículos, sino todo el Código, cuya extensión es de 26 páginas, ya que es importante que conozca, antes de iniciar el ejercicio de su profesión, las cuestiones éticas que reglamentan su actividad profesional.

Los artículos citados se relacionan con el respeto a la integridad de la persona que solicita el servicio psicológico, velar por su bienestar, no intervenir si no está en el ámbito de competencia del psicólogo, a quién hay que comunicar los resultados y en qué casos, considerar que las pruebas psicológicas son sólo instrumentos auxiliares, el derecho de los evaluados de conocer los resultados de la evaluación, utilizar sólo pruebas confiables y válidas, entre otros.

Es muy importante, al iniciar un proceso de evaluación psicológica, el consentimiento informado del usuario, explicarle porqué se le va a evaluar, cómo va a ser este proceso y pedirle su consentimiento para la aplicación de las diferentes pruebas, garantizándole la confidencialidad de la información proporcionada y obtenida y haciéndole saber que, excepto en casos excepcionales, los resultados se le entregarán a él. La ética del psicólogo puede resumirse dos grandes rubros: en el

---

respeto hacia la persona que solicita el servicio y en la procuración de su bienestar emocional.

## REFERENCIAS BIBLIOGRÁFICAS

- Aragón, L. y Silva, A. (2002) **Fundamentos teóricos de la evaluación psicológica**. México: Pax.
- Fernández, R. (1996) **Introducción a la evaluación psicológica**. Salamanca: Pirámide.
- Garaigordobil, M. (1998) **Evaluación Psicológica. Bases teórico-metodológicas, situación actual y directrices de futuro**. Salamanca: Amarú Ediciones.
- Martínez, R. (1996) **Teoría de los tests psicológicos y educativos**. Madrid: Síntesis Psicológica.
- Muñiz, J. (1992) **Teoría clásica de los tests**. Madrid: Pirámide.
- Pawlik, K. (1980) **Diagnosis del diagnóstico**. Barcelona: Herder.
- Sociedad Mexicana de Psicología (2001) **Código Ético del Psicólogo**. México: Trillas.