

CAPÍTULO 2

CONCEPTOS BÁSICOS PARA LA ASIGNATURA

2.1. Introducción

En este capítulo se estudiarán algunos conceptos básicos para esta asignatura como son la fiabilidad y validez de un tests, lo que se entiende por muestra tipificada y población, y las normas de tipificación, criterios estos indispensables para garantizar la objetividad de un tests y de su interpretación; algunas normas ya están implícitas en el propio proceso de construcción de los tests, pero existen otras normas tipificadas que se deben cumplir en el proceso de aplicación y calificación de los mismos, las que serán objeto de estudio en el capítulo siguiente.

Los resultados de los tests psicológicos constituyen datos objetivos de mayor o menor confianza en dependencia del cumplimiento de las normas establecidas para su construcción y elaboración (información que se obtiene a través del manual para el uso de dicho instrumento), pero también el grado de confiabilidad de los datos que se obtiene de los tests depende del desempeño del experto que aplica y califica estos instrumentos; de igual forma es importante conocer, para otorgar determinado grado de confianza a estos datos, el contexto cultural y personal del sujeto que se estudia.

La información resultante de la aplicación de instrumentos de evaluación que integran al proceso evaluativo, constituye indicadores valiosos ya que permite establecer criterios objetivos que posibilitan la formulación de hipótesis de trabajo; pero el resultado de los tests con el que se trabaja para la evaluación psicológica no siempre son las anotaciones brutas derivadas de la suma de la calificación de los ítems. Por ejemplo en los test de inteligencia que estudiaremos, una puntuación bruta no tiene ningún significado, como tampoco lo tiene una puntuación en porcentajes.

Decir que un sujeto resolvió el 30% de un test ofrece poca información acerca de esa persona.

El resultado obtenido en un test dice algo acerca de esa persona, cuando su calificación

representa un valor de ejecución del test, en una muestra tipificada que ha sido tomada de la población a la cual pertenece esa persona y tenido en cuenta, además, las variables que para dicho test tienen que ser contempladas.

2.2. Evaluación psicológica

La Evaluación Psicológica es una disciplina de la Psicología que se ocupa de la medida de algunos aspectos del comportamiento humano; esta medición se realiza por medio de metodologías básicas entre las que se encuentran las técnicas psicométricas y las proyectivas y otros métodos como la observación, la entrevista, los autoinformes, etc., que serán objeto de estudio en el segundo año. Todos estos instrumentos requieren rigurosos procesos para su selección y metódicos pasos en su aplicación y calificación.

El objetivo de la evaluación psicológica puede ser: Describir, Clasificar, Predecir, Controlar. Estas acciones se pueden definir como:

- a) Describir las características de personalidad de un sujeto.
- b) Clasificar es decir, enmarcar a la persona en una categoría determinada.
- c) Predecir el futuro comportamiento de un sujeto.
- d) Controlar determinadas características actuales en una persona para confrontarlas con posibles cambios de comportamiento en un momento posterior.

Estos objetivos de la evaluación psicológica son utilizados en la práctica del psicólogo para Diagnosticar, Seleccionar, Orientar, o Modificar comportamientos. Para ello el psicólogo debe poseer amplios conocimientos teóricos, empíricos y metodológicos básicos, que son fundamentales en la práctica de la Psicología. Dichos conocimientos contemplan los fundamentos biológicos y sociales del comportamiento, los procesos de la psicología general como son atención, pensamiento, percepción, memoria, aprendizaje, resolución de problemas, personalidad, las diferencias individuales y su evolución, conocimientos acerca de la cultura de la población a la cual pertenece la persona en estudio y conocimientos metodológicos que se requieren a la hora de la investigación en psicología, todo ello es necesario y es un conocimiento previo que debe poseer el psicólogo para realizar la evaluación psicológica en la

práctica del caso individual en lo que se refiere a la exploración y análisis del comportamiento de concretos seres humanos.

2.3. Evaluación psicológica y calificación de tests mentales

Es fácil entender que el resultado que arrojan estos instrumentos por sí solos no representa indicadores certeros para describir, predecir, valorar o diagnosticar. Los resultados de los instrumentos de evaluación psicológica se integran en un proceso de análisis del comportamiento humano, lo cual posibilita la toma de decisiones acerca de la persona estudiada; este análisis requiere del concurso de amplios conocimientos de la psicología que posibiliten una integración adecuada de dicha información.

La calificación de estos instrumentos permite obtener un resultado "objetivo" que posibilita al psicólogo tomar cierta distancia de sus juicios subjetivos acerca de la persona estudiada, conocer nueva información e integrar esa búsqueda al proceso de evaluación psicológica.

Los tests constituyen instrumentos de medida, el uso correcto o incorrecto de los mismos a la hora de efectuar las mediciones repercute en todo el proceso evaluativo y por ende, puede ser utilizada esta información correcta o incorrectamente, con el consecuente resultado que tal situación trae aparejada: atrasos en el trabajo, volver a calificar las pruebas y en ocasiones esos datos deben ser desechados y peor aún, ofrecer resultados erróneos acerca de una persona, partiendo de información que fue tomada imperfectamente.

2.4. Qué es un test

Según Anastasi, un test psicológico constituye esencialmente una medida objetiva y tipificada de una muestra de conducta. Podemos afirmar entonces que proporciona una medida objetiva, de lo que se intenta medir, cuando:

✂ Existe uniformidad de criterios, a ejecutar por todos los examinadores, en el proceso de aplicación del test en lo referente a consigna, condiciones generales de aplicación, como son:

iluminación, ventilación, preparación de los materiales necesarios, la forma de responder a las preguntas de las personas examinadas, y todos los detalles de la situación de examen. Esto se conoce como tipificación en el proceso de aplicación de un test determinado.

La tipificación incluye además, la velocidad con que se habla, las inflexiones de la voz, las pautas, la expresión facial, etc.

✂ Se mantiene el procedimiento tipificado de calificación, es decir, se otorgan las puntuaciones tal como está pautado para dicho test.

✂ La interpretación de las puntuaciones son objetivas en cuanto son independientes del juicio subjetivo del examinador en particular.

Es decir, el proceso de interpretación esta igualmente tipificado para las diferentes puntuaciones del test, en relación a la muestra de conducta que el test mide. En ocasiones la tipificación interpretativa de la puntuación incluye la consideración de algunas variables como la edad, el sexo, etc., ya que estas variables pueden introducir diferencias significativas para la muestra de conducta que se estudia.

✂ Existen normas tipificadas de la conducta que el test intenta medir para la población en la cual se aplica el test, lo que permite la interpretación de los resultados obtenidos por ese instrumento.

Para Cronbach, 1968, un test es objetivo cuando todos los examinadores asignan el mismo valor a una ejecución determinada. Si bien esto no se ha logrado del todo, hoy día todos los investigadores que intentan la construcción de test mentales se esfuerzan en el proceso de tipificación del instrumento siguiendo este objetivo. El uso de los test ha demostrado que los mismos constituyen un medio auxiliar, de inestimable valor, en el trabajo del psicólogo, aunque como ya se ha dicho nunca puede ofrecerse una explicación de la conducta humana desde la evaluación psicométrica.

2.5. Normas, población y muestra tipificada

La norma es la actuación media o normal del comportamiento en la población, de la variable o variables que el test mide. Sin normas, no es posible interpretar las puntuaciones del test. La puntuación de un individuo sólo puede valorarse comparándola con las obtenidas por otros. Por ejemplo, si los niños normales de 8 años ejecutan correctamente 12 de los 50 ítems de un determinado test de razonamiento aritmético, entonces la norma en este test, para los niños de 8 años de edad, corresponde a una puntuación de 12. La puntuación que obtiene una persona en un test no tiene ninguna significación hasta que se valora en función de un adecuado conjunto de normas.

Las normas de actuación para un test en una población determinada, se obtienen mediante el proceso de tipificación de dicho test en esa población.

Es necesario que el mismo se aplique a una gran muestra de personas, que represente el tipo de sujetos a los que se destina esa prueba.

Este grupo conocido como muestra de tipificación, sirve para establecer las normas que, no solo indican el promedio de actuación, sino también la frecuencia relativa de los diversos grados de desviación, por encima y por debajo del promedio. De esta forma, es posible valorar la posición del individuo, en cuanto a la muestra de conducta que se mide, respecto a la población a la cual pertenece.

Por población se entiende el grupo de personas de un determinado territorio que cumplen las variables que son tenidas en cuenta por dicho test. Eso quiere decir que si un test está elaborado para ser aplicado en niños de 6 a 11 años de edad, y se desean obtener los valores normativos para dicha prueba, en los niños de estas edades, en una ciudad determinada; la población con la cual se tiene que trabajar para obtener los valores normativos para los niños de esas edades, para ese test en particular, tienen que ser todos los niños que pertenecen a dicha ciudad y están comprendidos entre las edades de 6 a 11 años. En la práctica, no se aplica el test a toda esa población, utilizando la Estadística se pueden obtener los valores normativos trabajando tan solo con una muestra de los niños de 6 a 11 años de edad de la ciudad; ese

grupo de niños es lo que se conoce como muestra de tipificación.

Veamos a continuación cómo Anastasi explica uno de los procedimientos que se sigue para que un test sea considerado como objetivo; este procedimiento está basado en el grado de dificultad que alcanzan los ítems del test o el test completo, es lo que se conoce como Medida objetiva de dificultad, y en base a esta medida de dificultad es que se puede realizar el trabajo de diferenciación de unas personas en relación a otras y su ubicación, de acuerdo a la distribución en una curva de normalidad, del rasgo que se mide con dicho test.

“Cuando Binet y Simon prepararon su escala original, la de 1905, para la medida de la inteligencia (cf. Cap. I), dispusieron los 30 elementos de la escala en orden de dificultad creciente. Se recordará que esta dificultad se determinó ensayando los elementos en 50 niños normales y unos pocos retrasados y débiles mentales. Los elementos resueltos correctamente por el mayor número de sujetos se consideraron ipso facto como los más fáciles; aquellos que acertaron relativamente pocos sujetos se consideraron como más difíciles. Mediante dicho procedimiento se estableció un orden empírico de dificultad. Este antiguo ejemplo tipifica la medida objetiva del nivel de dificultad, que ahora se considera como práctica común en la elaboración de un test psicológico.

No solo la disposición, sino también la selección de los elementos para su inclusión en un test, pueden determinarse por la proporción de sujetos que pasan cada elemento en las muestras de prueba. De esta forma, si hay una reunión de elementos en el extremo fácil o difícil de la escala, algunos de ellos pueden descartarse. De la misma manera, si los elementos escasean en ciertas porciones del campo de dificultad, cabe añadir nuevos elementos para rellenar los huecos.

La frecuencia de respuestas correctas se emplea también en la elaboración de las escalas de edad, como sucede en las últimas revisiones de las escalas de Binet. En este caso se determina la proporción de niños que pasan cada elemento en cada nivel de edad. Se asigna entonces el elemento a aquel nivel de edad en el que lo acierta una proporción determinada.

El nivel de dificultad del test, considerado en su conjunto, depende directamente, desde luego,

de la dificultad de los elementos que lo constituyen.

La distribución de las puntuaciones totales proporciona un control completo de la dificultad del test total para la población a la cual se destina. Si la muestra de tipificación es una sección transversal representativa de tal población, entonces suele suceder que las puntuaciones figuran aproximadamente dentro de una curva de distribución normal. En otras palabras, debe haber una acumulación de individuos cerca del centro del campo de variabilidad y una disminución gradual del número de los mismos a medida que nos acercamos a los extremos. En la figura siguiente se representa una curva normal teórica, con todas las irregularidades eliminadas. Al representar gráficamente esta distribución de frecuencias, se indican las puntuaciones en la base, y las frecuencias, o número de personas que obtienen cada puntuación, en el eje vertical. Cuando se aplican tests a muestras muy grandes se obtiene una curva suave muy aproximada a la de la figura 2.1.

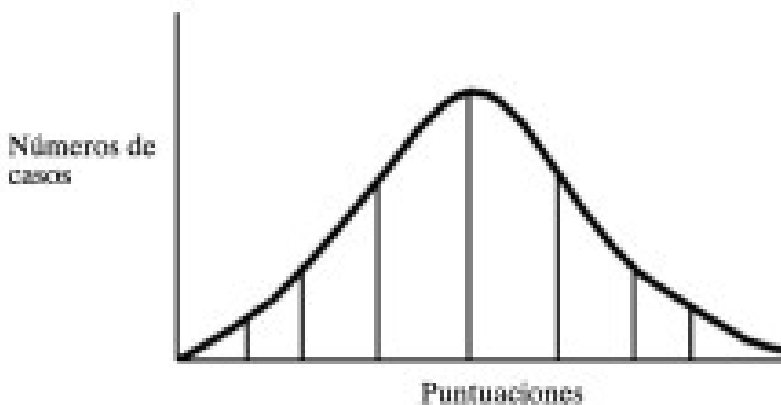


Fig. 2.1. Curva de distribución normal.

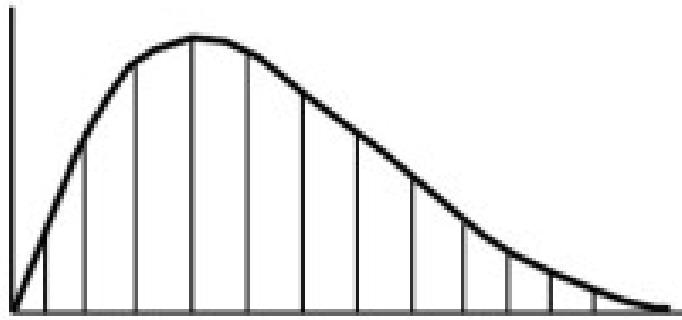
Supongamos, sin embargo, que la curva de distribución obtenida no es normal, sino claramente desviada, o torcida, como las que presentamos en la figura 2.2. La primera, con una acumulación de puntuaciones en el extremo inferior, sugiere que el test tiene un suelo demasiado alto para el grupo que se considera, faltándole un número suficiente de elementos fáciles para discriminar adecuadamente en el extremo inferior del campo. El resultado es que las personas que normalmente se esparcirían en amplitud considerable obtienen en este test puntuaciones de cero o próximas a cero. Por consiguiente, se obtiene una cúspide en el extremo inferior de la

escala. En la figura 2.3.se muestra esquemáticamente dicha acumulación artificial de puntuaciones, en la cual un grupo distribuido normalmente produce una distribución desviada en un test determinado.

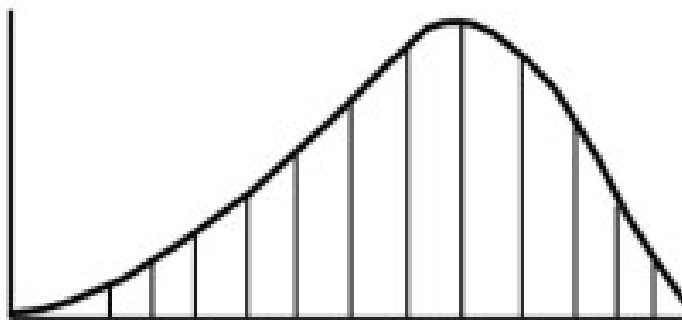
En la figura 2.2.B tenemos un ejemplo de una desviación opuesta, con las puntuaciones acumuladas en el extremo superior, resultado que sugiere insuficiencia de techo en el test. Cuando un test destinado a la población en general se aplica a muestras seleccionadas de estudiantes de enseñanza superior o graduados, se suele producir corrientemente una distribución desviada, ya que algunos estudiantes obtienen puntuaciones casi perfectas. Con un test de este tipo es imposible medir las diferencias individuales entre los sujetos más capacitados del grupo. Si se hubieran incluido elementos más difíciles, algunos individuos habrían puntuado más alto de lo que les permite este test.

Cuando en un test la muestra de tipificación arroja una distribución claramente fuera de lo normal, suele modificarse el nivel de dificultad hasta que se aproxima a la curva normal. De acuerdo con el tipo de desviación de la normalidad que aparezca, pueden añadirse elementos más fáciles o más difíciles, eliminar o modificar otros elementos, alterar la posición de los mismos en la escala o revisar los pesos o ponderaciones de la puntuación asignada a ciertas respuestas. Se continúa con estas adaptaciones hasta que la distribución llega a ser normal, al menos aproximadamente.

En estas condiciones, la puntuación más probable obtenida por el mayor número de sujetos suele corresponder aproximadamente al 50 por 100 de elementos correctos. Para el lego en la materia, que no se halla familiarizado con los métodos de elaboración de un test psicológico, una puntuación de un "50 por 100" puede parecerle extraordinariamente baja. A veces se objeta, partiendo de esta base, que el examinador ha asignado un techo demasiado bajo al test, o se supone que el grupo al que se ha aplicado es particularmente pobre. Estas dos conclusiones carecen de valor cuando se contemplan a la luz de los procedimientos que se emplean en la elaboración de los tests psicológicos.



a. Acumulación en el extremo inferior de la escala



b. Acumulación en el extremo superior de la escala

Fig. 2. 2. Curvas de distribución desviadas

Estos se elaboran deliberadamente y se modifican específicamente, de forma que arrojen una puntuación media de un 50 por 100 de respuestas correctas, aproximadamente.

Solo de esta manera puede obtenerse con el test el máximo de diferenciación entre los individuos en todos los niveles de aptitud.

Con una media de un 50 por 100 aproximadamente de elementos correctos existe la máxima posibilidad de obtener una distribución normal, con puntuaciones individuales que se extiendan ampliamente por ambos extremos.

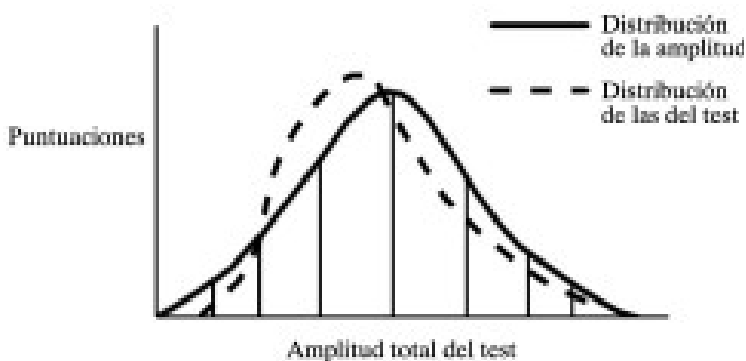


Fig. 2.3. Desviación resultante por insuficiencia del suelo en el "test"

2.6. Validez y fiabilidad

Además de las tipificaciones antes mencionadas, un test debe poseer validez y fiabilidad para considerarse un instrumento objetivo de medición.

Estos dos factores están estrechamente ligados, ya que un instrumento no puede ser válido si no es al mismo tiempo fiable. Examinemos algunas modalidades en relación a la validez y fiabilidad, descritos por Anastasi.

Fiabilidad significa la estabilidad o consistencia de la puntuación, obtenida por una persona, en diferentes momentos en que se le aplique el mismo test. Por ejemplo, si una persona obtiene un CI de 110 en una ocasión y días después 80 al repetir el test, es obvio que esas puntuaciones ofrecen poca confianza y por ende, el instrumento de medida utilizado carece de fiabilidad. Cuando se dice que un test es fiable se supone que su diseñador ha tenido en cuenta todos los indicadores que aseguran tal aseveración. Citamos algunos de los aspectos a tener en cuenta en la fiabilidad de un test:

Estabilidad temporal: en el ejemplo anterior la fiabilidad del test utilizado para medir el CI se ve afectada por lo que se denomina estabilidad temporal.

La estabilidad temporal indica el grado en que las puntuaciones de un test quedan afectadas por las fluctuaciones diarias que se producen en el sujeto o en el ambiente en que se aplica el test. La estabilidad temporal de un test depende parcialmente de la longitud del intervalo sobre el que se mide. Algunos tests muestran elevada fiabilidad tras períodos cortos, pero cuando el intervalo de aplicación se extiende en el tiempo se observa una falta casi absoluta de correspondencia entre las puntuaciones obtenidas en los dos momentos diferentes. Ejemplo de ello son algunos test que miden CI en etapa pre-escolar, que arrojan medida de estabilidad moderada dentro del período preescolar pero que no permiten hacer predicciones de los CI del último período de la niñez o de otras etapas posteriores de la vida.

Para cualquier tipo de sujeto, el intervalo de repetición de los tests rara vez debe pasar de los 6 meses. En intervalos de tiempo largos es muy posible que las condiciones de vida en las

cuales se desempeña un sujeto cambien, al punto tal que su rendimiento en los test puede mejorar o empeorar, en relación a él mismo o en comparación con sujetos de su misma edad, a causa de circunstancias como pueden ser: desarrollo de habilidades por algún oficio, condiciones de su hogar o comunidad o por otras razones como una enfermedad o trastorno emocional. Factores de este tipo hacen necesario realizar nuevas evaluaciones en personas que ya fueron evaluadas con anterioridad.

Fiabilidad del examinador para la aplicación y la calificación. En este sentido casi todos los tests proporcionan procedimientos tipificados para su aplicación y puntuación que permiten pensar que la fiabilidad del examinador para la aplicación y calificación es alta a los fines prácticos.

Lo único a vigilar es asegurarse de que se sigue con cuidado los procedimientos prescritos. El problema reside en el control empírico de las condiciones bajo las cuales el test debe ser aplicado.

La validez se refiere al grado en que el test mide lo que pretende medir. De igual forma existen varios procedimientos para determinar la validez de un test:

Validez de contenido. Se refiere al examen sistemático del contenido del test para determinar si comprende una muestra representativa de la forma de conducta que ha de medirse.

Validez predictiva. Indica la efectividad del test en la predicción de algún resultado futuro.

Validez concurrente. La relación entre las puntuaciones del test y los índices de status del criterio obtenido, aproximadamente al mismo tiempo, se conoce como validez concurrente. Ejemplo, las puntuaciones altas obtenidas en un test de inteligencia por un grupo de estudiantes universitarios se contrastan con los buenos resultados académicos obtenidos por estos. Esta convergencia entre los resultados del test y la acumulación de notas ofrecen lo que se llama validez concurrente.

Validez de elaboración o de hipótesis de trabajo. La validez de elaboración de un test es el

grado en que este mide una elaboración teórica o rasgo.