

Nombre del alumno:

Sofia Gómez Avila

Nombre del profesor:

Liliana Villegas López

Nombre del trabajo:

Resumen “Confiability”

Materia:

Psicometría

PASIÓN POR EDUCAR

Grado:

Tercer cuatrimestre

Grupo: u

CONFIABILIDAD

Cuando se habla de confiabilidad en psicometría se refiere a la consistencia de resultados, es decir que sea coherente y que dure. Las pruebas son para los psicólogos una herramienta para obtener información, lo cual, a su vez, les permite plantear, luego, una impresión diagnóstica con miras a desarrollar una acción preventiva o de intervención, según la dinámica del caso.

Las pruebas psicológicas, nos permiten medir atributos humanos y son las herramientas más cercanas para los psicólogos cuando de medir estos atributos se trata. Pero, al mismo tiempo, debemos tener claro que no son las únicas y que por sí solas no pueden lograr la confiabilidad de un proceso evaluativo.

Actualmente, coexisten dos teorías generales de los tests, la teoría clásica de los tests y la de respuesta al ítem.

La hipótesis fundamental de la teoría clásica de los tests (TCT) es que la puntuación observada de una persona en un test es una función de dos componentes: su puntaje verdadero (que es inobservable) y el error de medición implícito en toda medición. La puntuación observada de un individuo es igual a la puntuación verdadera más el error. En el plano teórico, la puntuación verdadera puede entenderse como la media de las puntuaciones obtenidas por una persona en infinitas aplicaciones de un test (en diferentes momentos y condiciones), asumiendo que la forma de distribución de esas infinitas puntuaciones se aproxima a la normal.

La puntuación de error, por otra parte, es la suma de todos aquellos factores aleatorios que influyen y afectan el registro de los datos, introduciendo inconsistencia en el proceso y alejando la puntuación observada de la puntuación verdadera. La confiabilidad puede entenderse como la exactitud o precisión de una medición, o el grado en el cual las puntuaciones de un test están libres de esos errores de medición. Esta exactitud o precisión de las puntuaciones permite que éstas se mantengan constantes en diferentes circunstancias.

Es importante diferenciar entre **errores sistemáticos y aleatorios**. Un error sistemático es aquel que genera constantemente una puntuación elevada o baja en una persona al margen de los cambios que puedan darse en el rasgo medido por el test. Este tipo de error se denomina sesgo de medición, que se define como toda aquella fuente de variación que sistemáticamente afecta a las puntuaciones de un determinado grupo que está siendo evaluado por un test, ya sea elevando o disminuyendo las mismas. Un instrumento que mide aptitudes para la matemática puede estar construido de tal manera que sus ítems sean poco comprensibles para un grupo social o etario determinado, y esto podría ocasionar que ese grupo

obtenga puntuaciones más bajas en el test independientemente de sus aptitudes matemáticas.

En general, la confiabilidad se verifica mediante un coeficiente de correlación entre las medidas repetidas de un fenómeno. Para entender la lógica de la utilización de este coeficiente cabe realizar algunas observaciones previas. La puntuación verdadera también puede expresarse en términos de varianza de las puntuaciones de los tests. Recuérdese que la varianza indica la dispersión promedio de los valores (en este caso los puntajes de tests) alrededor de la media de un grupo de observaciones (más precisamente, el promedio de los cuadrados de la diferencia entre cada valor y la media). En la teoría clásica de los tests, la confiabilidad puede expresarse como la proporción de varianza observada de las puntuaciones de tests que se debe a la varianza verdadera (la variabilidad de la puntuación verdadera), es decir, a la varianza del rasgo que se está evaluando y no a otros factores.

Los diferentes métodos utilizados para evaluar la confiabilidad utilizan el coeficiente de correlación como estadístico fundamental.

La confiabilidad (como la validez) no es una característica del test en sí mismo, según la teoría clásica de los tests, sino una propiedad de las puntuaciones del test cuando éste se administra a una muestra específica y bajo condiciones particulares.

Por otro lado, existen múltiples factores que afectan la confiabilidad de las puntuaciones obtenidas mediante la aplicación de un test. La APA (1999) los clasifica en factores internos (fluctuaciones en el nivel de atención del examinado, por ejemplo) o externos (diferentes criterios de corrección de un test, según evaluadores distintos, por ejemplo).

Como primer punto, tenemos el *contenido del test*. Al construir un test debe tenerse en cuenta que la elección de los ítems, los materiales a través de los cuales estimulamos la respuesta del individuo, y la formulación de las consignas pueden ser una fuente de error aleatorio en la medición. Si se desea estimar el puntaje de un individuo en un cierto dominio, los errores en el muestreo de contenido pueden afectar la confiabilidad.

Si consideramos a un test como un conjunto de indicadores operacionales de un rasgo o dominio de comportamiento, es comprensible que la elección de los mismos (muestra de ítems) pueda constituir una fuente de error. Tanto si el muestreo de contenido es sesgado, como si no es suficientemente extenso, los puntajes resultantes serán poco confiables. Un test de inteligencia con solo 10 ítems representa un ejemplo de muestreo de contenido muy limitado, pues en este caso los puntajes dependerían de la capacidad del individuo con respecto al pequeño conjunto de reactivos utilizados. Si las puntuaciones de un test son

afectadas por el muestreo de contenido o por el uso de una determinada escala de respuesta, las diferencias que se observen en las puntuaciones de las personas evaluadas se deberán no a diferencias reales en el rasgo medido sino a variaciones relacionadas con errores de medición.

Otro de los factores es la *administración*, en el momento de administrar un test también pueden introducirse errores que afecten la confiabilidad de los resultados. Por ese motivo es esencial examinar a todos los participantes en condiciones uniformes, estandarizadas. Así, por ejemplo, las condiciones generales del ambiente en que se administran los tests deben ser lo más semejantes que sea posible.

Por consiguiente, siempre es deseable que las instrucciones del test sean lo suficientemente claras y unívocas, para que todos los evaluadores las impartan de la misma manera y presenten los materiales en idéntico orden y forma a todos los examinados. La falta de consistencia en la administración de un test influirá en la estabilidad de las puntuaciones obtenidas por las personas medidas por ese test.

La *calificación*. En el momento de calificar (puntuar) un test existen otros factores que pueden influir negativamente en la confiabilidad de las puntuaciones obtenidas. En los tests de opción múltiple pueden cometerse errores cuando la corrección es manual, tales como calificar accidentalmente algunas respuestas correctas como erróneas y viceversa, o realizar mal la sumatoria de las respuestas clave o correctas. La posibilidad de cometer este tipo de errores prácticamente se elimina en los tests computarizados. Cuando en el proceso de puntuación de un test interviene de manera importante el criterio del evaluador, puedan presentarse variaciones que disminuyan la confiabilidad de los puntajes.

Factores internos del examinado. Parte de la varianza atribuible a errores de medición se origina en las fluctuaciones azarosas del comportamiento de la persona examinada, que aumentan o disminuyen su puntaje. En este sentido, las distracciones momentáneas, las preocupaciones de índole personal y otros acontecimientos semejantes pueden afectar la estabilidad de los puntajes de tests. Aunque tales influencias fortuitas tienden a atenuarse en una prueba extensa y bien estandarizada, las variaciones en la atención o motivación de las personas al realizar el test siempre pueden afectar su confiabilidad.

La confiabilidad también varía en función de la muestra utilizada para estimarla, lo cual constituye una de las limitaciones más serias del modelo clásico de los tests. La confiabilidad aumenta al incrementarse la variabilidad de las respuestas y, por consiguiente, se recomienda que las muestras empleadas para verificar la confiabilidad de un test sean lo más heterogéneas posible en aquellas características que generan diferencias entre las personas que la integran (sexo,

nivel socioeconómico, nivel educativo, por ejemplo), a los fines de incrementar la variabilidad de las respuestas.

En conclusión, es muy importante tomar en cuenta los resultados de tests ya que si no existe un resultado concreto o confiable afecta, tanto la integridad del paciente como del psicólogo.

Finalmente, vale la pena decir que el trabajo del psicólogo dedicado a la psicometría y evaluación en Psicología es de gran valor y utilidad. Requiere, en síntesis, un buen manejo de la estadística, mucha creatividad para diseñar tests, rigurosidad científica, manejo ético y una mente muy disciplinada.