
GLOSARIO DE TÉRMINOS PSICOMÉTRICOS Y AREAS AFINES

Nelson Rodríguez Trujillo Ph.D.

www.psycoconsult.com

psycoconsult@cantv.net

GLOSARIO DE TÉRMINOS PSICOMÉTRICOS Y AREAS AFINES

Recopilado y elaborado por Nelson Rodríguez Trujillo Ph.D.
Escuela de Psicología
Universidad Central de Venezuela
Psico Consult C.A.
Versión junio, 1999

Este Glosario de Términos Psicométricos y Áreas Afines se desarrolló para apoyar el aprendizaje de los alumnos de la Cátedra de Psicometría de la Escuela de Psicología de la Universidad Central de Venezuela. La recopilación se inició en 1994 con 52 términos incluidos en el Programa PsicoMet (desarrollado por Psico Consult C.A. y Sistemas L+D), software que permite la corrección automatizada de pruebas y la elaboración de informes psicométricos; también fue incluido en el curso Pruebas Psicométricas y Decisión de Selección. Se ha ido enriqueciendo sistemáticamente con aportes provenientes de diferentes fuentes, entre las cuales merecen mención particular los alumnos de Psicometría II de 1998, y ahora incluye 120 términos. Las definiciones fueron elaboradas haciendo uso de conceptos propios y de diversas fuentes bibliográficas, algunas de las cuales se mencionan al final. Dado que siempre habrá términos adicionales y mejores definiciones que agregar, este Glosario puede considerarse en proceso de mejora constante, por lo que se agradecen las contribuciones que puedan dar sus usuarios.

Los términos expresados en *letras Itálicas* refieren a otros términos contenidos en el Glosario como una ayuda para ampliar las definiciones.

Actitud y Evaluación de Actitudes. Se entiende por actitud una predisposición organizada y duradera con componentes cognoscitivos, afectivos y conductuales, que orienta a la persona a pensar, sentir, percibir o a comportarse de una manera determinada hacia un referente u objeto cognoscitivo. Se evalúan usualmente mediante inventarios, es decir, listas de afirmaciones asociadas con el objeto de la actitud, a las cuales el *Sujeto* responde, indicando en una escala su disposición positiva o negativa hacia esa afirmación.

Análisis de ítems. Procedimiento cuantitativo mediante el cual se determina el funcionamiento de los *Ítems* de una prueba o encuesta. Permite establecer el número y proporción de sujetos que responde la *Clave* y cada uno de los *Distractores*, así como la *Correlación Biserial* de las opciones con el puntaje total en la prueba o con un *Criterio de Validación* externo.

Análisis Factorial. Véase *Factor*

Anava de Hoyt. Procedimiento para el cálculo de la *Confiabilidad* de Consistencia Interna de una prueba, que establece la relación entre la varianza debida a los sujetos, los ítems y la varianza total del test.

Aptitud y Pruebas de Aptitudes. Una aptitud es una combinación de habilidades, conocimientos y otras características, innatas o adquiridas, que indican el potencial de un individuo para aprender a desarrollar un buen rendimiento en un área particular si se le provee de educación y entrenamiento. Entre los tests aptitudinales se incluyen los que evalúan la habilidad académica general (llamados comúnmente de habilidad mental o de inteligencia), los de habilidades especiales tales como verbal, numérica, mecánica o musical; tests de apresto para el aprendizaje; tests de pronóstico que evalúan tanto la habilidad como el aprendizaje anterior y son usados para predecir el rendimiento futuro – usualmente en un campo específico, como una lengua extranjera, taquigrafía, etc. También se habla de “aptitud” en un sentido más amplio, por lo que la “aptitud musical”, por ejemplo, se referiría a una combinación de características físicas y mentales, y de aspectos motivacionales, intereses y otras características, que conducen a lograr un buen rendimiento en música.

Atenuación de la correlación Véase *Corrección por Atenuación*.

Batería de Pruebas. Una Batería de Pruebas es un grupo de tests estandarizados sobre una misma muestra de población, de manera que los resultados en los diferentes tests son comparables. También se concibe como un conjunto de tests desarrollados para evaluar habilidades interrelacionadas entre sí teóricamente y/o empíricamente y asociadas con la realización de una tarea o con los estudios académicos. Cuando se aplican todas las pruebas como batería, es posible elaborar un *Perfil Psicométrico* de la persona. Es usual que las variables a explorar con los tests de una Batería se seleccionen de manera que tengan validez factorial o que tengan altas correlaciones entre los ítems de una prueba, pero bajas entre pruebas. Un ejemplo de batería es el Test Diferencial de Aptitudes de Bennet y Seashore.

CEEB. Transformación lineal derivada introducida por el College Entrance Examination Board de USA, de Media 500 y Desviación Típica 100. Su gran

amplitud permite la comparación de resultados de grandes grupos de sujetos, acumulados de año en año.

Clave de Corrección. Es el conjunto de respuestas correctas a las preguntas de una Prueba Psicométrica. La clave de corrección, también debe indicar la ponderación dada a cada una de las respuestas para el cálculo del puntaje bruto total.

Cociente Intelectual (CI). Indicador del nivel de inteligencia utilizado en las primeras pruebas psicométricas y que se calculaba mediante la división de Edad Mental (calculada por el test) y Edad Cronológica (edad real del *Sujeto*) multiplicada por 100. Un valor de 100 (igual edad mental y edad cronológica) indica normalidad, mientras que por debajo indica "retardo" y por encima "aceleración".

Cociente Intelectual de Desviación (CI Desviación). Sustituto del *Cociente Intelectual* original se desarrolló por las diferencias existentes en la varianza de las puntuaciones del Cociente Intelectual en los diferentes niveles de edad, que dificultaban la interpretación de los resultados. Es en realidad una transformación que tiene una media de 100 y una desviación típica de 15.

Colectivas, Pruebas. Pruebas que se pueden aplicar a grupos de sujetos para su evaluación. Este concepto se contrapone al de Pruebas Individuales.

Coefficiente Alfa. Estimado de confiabilidad de consistencia interna propuesto por Cronbach. Para ítems binarios, da resultados iguales a la fórmula Kuder Richardson 21 (*KR21*).

Coefficiente de Correlación. Medida del grado de asociación que existe entre dos conjuntos de medidas realizadas en el mismo grupo de individuos. El más conocido y utilizado con Pruebas Psicométricas es *Producto Momento de Pearson*, aunque existen otros como correlación por rangos, biserial, tetracórico, etc. Tiene un recorrido entre +1.00 y -1.00 (asociación perfecta positiva o negativa), pasando por 0.00 (ausencia de relación entre ambas medidas). La correlación indica solamente asociación entre variables, no relaciones causales.

Competencias. Características de las personas directamente relacionadas con el desempeño exitoso de un trabajo o una actividad. Incluyen conocimientos, habilidades, disposiciones afectivas, rasgos físicos y conductas de carácter relativamente estables que se pueden aplicar a diferentes situaciones y tipos de trabajo o tareas.

Confiabilidad. El grado en que un test es consistente al medir la variable que mide. Algunas palabras asociadas con la Confiabilidad son estabilidad, exactitud, consistencia y precisión, así como el *Error Estándar de Medida*. La Confiabilidad se expresa usualmente mediante coeficientes de correlación, obtenidos a partir de los puntajes en dos aplicaciones de la misma prueba al mismo grupo de *Sujetos* o haciendo uso de pruebas paralelas aplicadas al mismo grupo de sujetos. También se define como Consistencia Interna, que se obtiene mediante la determinación de las intercorrelaciones o *Covarianzas* de los ítems (véase *KR21* y *Coefficiente Alfa*).

Consistencia Interna. Véase *Confiabilidad*

Construcción de pruebas, fases de la. Las pruebas psicométricas se construyen mediante un cuidadoso proceso, que garantiza la calidad de las mismas. Tavella identifica varias fases: 1) Preparatoria, en la que se define (teórica y operacionalmente) la variable a explorar, se establecen las especificaciones del instrumento, se elaboran los ítems, se depuran cualitativamente mediante el juicio de expertos, se elaboran instrucciones y se estructura el instrumento. 2) Exploratoria, en la cual se aplica a una muestra informal de sujetos de la misma población a la cual va dirigida, se realiza análisis de ítems para determinar su calidad y se depura el instrumento. 3) Experimental, en la cual se aplica el instrumento ya depurado en la fase anterior a muestras formales de sujetos, hasta conocer mejor el instrumento, elaborar estudios de validez y confiabilidad y normas. 4) Definitiva, en la cual ya se considera que el instrumento está listo para su uso en grupos extensos de la población, así como para la realización de estudios normativos a gran escala.

Constructo. Las variables psicológicas no son observables directamente, por lo que se infieren a través de la observación de la conducta de las personas, y se plantean en términos teóricos. Son inventados y adoptados en forma deliberada y consciente por la comunidad científica, es decir, "construidos" por los investigadores y teóricos de la psicología.

Contaminación del Criterio. Situación en la cual, una persona que debe establecer un *Criterio de Validación* conoce de antemano las puntuaciones obtenidas por los sujetos en la prueba y sesga su juicio por ese conocimiento. Por ejemplo, un maestro que debe indicar el nivel de habilidad verbal de un grupo de alumnos, pero conoce con antelación los resultados de la prueba de razonamiento verbal que va a ser validada con su evaluación.

Corrección (puntuación) de Pruebas. Es la asignación de puntajes a las respuestas dadas por un sujeto a las preguntas de una prueba. La mayoría de las veces se realiza mediante la asignación de un punto por cada respuesta correcta (ítems binarios), es decir, que coincida con la clave de respuestas (ver Clave de Corrección), aunque también son posibles otras formas de corrección, como restar a las respuestas correctas una proporción de las incorrectas. Consulte también Puntaje Directo o Puntaje Bruto.

Corrección de Puntajes por Azar. Los ítems de opciones múltiples permiten responder al azar y alcanzar un cierto número de aciertos, es decir, permiten la adivinación de las respuestas. Por ejemplo, en una prueba de 40 ítems de cuatro opciones, la probabilidad de marcar la respuesta correcta por azar es de 25%, lo que permite establecer una media de puntajes por azar de 10 puntos y una desviación típica de 2,73 (en una distribución binomial, la media es $0,25 \times 40 = 10$ puntos y la desviación típica raíz de $0,25 \times 0,75 \times 40 = 2,73$). Por ello se han propuesto formas de corrección por puntajes por azar. La más común es la siguiente, en donde X_c es puntaje corregido, C es el número de respuestas correctas, I es el número de incorrectas y k es el número de opciones. Estas fórmulas asumen que toda respuesta incorrecta es una respuesta al azar; lo usual es que se le recomiende al *Sujeto*, que se abstenga de dar respuestas en las que

no esté completamente seguro y se le advierta sobre el uso de la fórmula de corrección.

$$X_c = C - \frac{I}{k-1}$$

Corrección por atenuación. La correlación entre dos variables se modera en su magnitud, por el nivel de confiabilidad de cada uno de las medidas. Esto hace que el resultado aparente ser más bajo, que la correlación que se obtendría si se contase con medidas perfectamente confiables. Para determinar el nivel de asociación que se obtendría con medidas confiables, se hace uso de la corrección por atenuación, que se obtiene por la fórmula siguiente, en donde r_{TG} es la correlación entre los puntajes verdaderos, r_{tg} es la correlación obtenida entre las dos variables, r_{tt} es la confiabilidad de una de las medidas y r_{gg} es la confiabilidad de la otra.

$$r_{TG} = \frac{r_{tg}}{\sqrt{r_{gg} \cdot r_{tt}}}$$

Correlación Punto Biserial. Estimado de la correlación Producto Momento de Pearson que se calcula entre una variable continua y una variable dicotómica. Se aplica en el Análisis de Items para determinar la asociación entre el puntaje total en el test y el puntaje obtenido en la Clave y cada uno de los distractores. Se calcula mediante la fórmula siguiente, en donde M_p es la media obtenida por las personas que respondieron correctamente el ítem, M_q la media de los que respondieron de manera incorrecta, p es la probabilidad de respuesta correcta del ítem, q la probabilidad de respuestas incorrectas y S_x la desviación típica de la prueba.

$$r_{pb} = \frac{M_p - M_q}{S_x} \sqrt{pq}$$

Correlación Múltiple. La correlación que se establece entre una variable predicha y varias variables predictoras en contraste con la *Correlación simple*. Por ejemplo, el rendimiento universitario (Criterio) que se predice haciendo uso de los resultados en varios tests (variables predictoras). La Correlación Múltiple tiene especial aplicación en las Ecuaciones de Regresión Múltiple, que permiten establecer una recta de regresión que incluye varias variables predictoras de la forma siguiente (en donde K es una constante, $\beta_n X_n$ es la ponderación que se le da a cada una de las variables predictoras):

$$Y' = K + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

Correlación Múltiple por Pasos. *Correlación Múltiple* que se establece entre una variable predicha y varias variables predictoras en forma progresiva. Para ello, se toma el coeficiente de correlación más alto de todas las predictoras con la

predicha, y se establece una recta de regresión, que permite establecer la correlación entre Y y Y'. La siguiente variable predictor se correlaciona con los residuos para establecer si correlaciona en forma significativa. Si es así, se incorpora al cálculo de la recta de regresión del siguiente nivel. Se deja de incorporar variables cuando una variable predictor no correlaciona en forma significativa con los residuos.

Correlación simple. Relación que se establece entre una variable predictor (o independiente) y una variable predicha (o Criterio). Es el modelo más sencillo de correlación. Véase *Correlación Producto Momento de Pearson*.

Correlación producto momento de Pearson. Coeficiente de correlación que permite establecer el grado de asociación entre dos variables continuas. Dado que se fundamenta en los mínimos cuadrados, exige que las variables guarden una relación lineal entre ambas y que presenten homocedasticidad. Se calcula mediante la fórmula:

$$r_{xy} = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{NS_x S_y}$$

Covarianza. Grado en que dos o más variables varían unas con relación a las otras. Pueden ser variables continuas o dicotómicas, como en el caso de los ítems. Se puede establecer mediante índices de covarianza, que en el caso de los ítems se determina mediante la fórmula siguiente: $cov. = p_{ij} - p_i \cdot p_j$. En realidad, la covarianza es la correlación entre dos variables multiplicada por sus desviaciones típicas.

Criterio de Validación. Es un estándar mediante el cual se evalúa la validez de un test. Se requiere que sean también una medida de algún tipo, es decir, que se expresen en una Escala Cuantitativa. Pueden ser el rendimiento en otra variable como otro test, la apreciación de un supervisor en el trabajo, los juicios de un profesor, notas en un curso, etc. Usualmente se establece una correlación entre el rendimiento en el test y la puntuación en el criterio para decidir cuál es el grado de asociación entre el test y el criterio, lo que permite predecir el rendimiento a futuro o determinar el nivel actual en el criterio.

Criterio, Pruebas referidas a. Véase *Pruebas referidas a Dominio*

Curva Normal. Véase *Distribución Normal*.

Desviación Típica. Es una medida de la variabilidad o dispersión de una distribución de puntajes. Mientras menos dispersos están las puntuaciones alrededor de la Media menor es la desviación típica, mientras más se dispersan, mayor es la Desviación Típica. Cuando se trata de una Distribución Normal, existen algunas relaciones importantes. Por ejemplo, a una desviación típica por encima y por debajo de la Media se encuentra el 68,26% de la distribución. El cálculo de la Desviación Típica se realiza mediante el cálculo de la raíz cuadrada de la sumatoria de los desvíos de los puntajes con respecto a la media elevados al cuadrado. Es también la raíz cuadrada de la *Varianza*.

Dificultad de los ítems. El grado de dificultad de un ítem se determina mediante la proporción de respuestas correctas (cociente entre respuestas correctas divididas entre el número de respuestas totales). La proporción que se obtiene se denomina a la media del ítem, o "la p del ítem" y es un indicador directo de su facilidad, o inverso de la dificultad, ya que su valor, comprendido entre 0 y 1, indica que mientras más se acerca a 0 es más difícil el ítem (puesto que fue respondido por un menor número de personas) y mientras más se acerca a 1 es más fácil (respondido por un mayor número de personas).

Distractores. Véase *Items*

Distribución bivariada normal. Modelo matemático que permite describir la distribución de los puntajes de un grupo en dos variables simultáneamente. Para utilizar este modelo, es importante que los datos que se describen con esta distribución muestren normalidad en sus distribuciones marginales, linealidad en su relación e igualdad de varianzas en todos los niveles de ambas variables (homocedasticidad).

Distribución Normal. Es una distribución de puntajes que tienen en su representación gráfica, la típica forma de una campana. Es llamada también Campana de Gauss. En una distribución normal, los puntajes se distribuyen en forma simétrica alrededor de la media, con tantos casos hacia arriba como hacia abajo para distancias iguales. Los casos se concentran cerca de la media y disminuyen en frecuencia a medida que se alejan de la media, de acuerdo a una ecuación matemática muy precisa. La Media, Mediana y Moda son idénticos en una Distribución Normal, ya que existe simetría. No toda distribución simétrica es necesariamente normal. Nada indica en realidad que los rasgos psicológicos se distribuyan de manera normal en la población de manera independiente del test que los mide; sin embargo, asumir que se distribuyen en forma normal en la población o en muestras de la misma y hacer que las puntuaciones en el test se distribuyan en forma normal, es útil porque permite hacer uso de las propiedades conocidas de la Curva Normal para obtener conclusiones y realizar predicciones.

Efecto Halo. Efecto subjetivo que ocurre cuando al evaluar a una persona la impresión que causa una variable o aspecto del sujeto afecta la evaluación en otras variables. Por ejemplo, la buena presencia de una persona determina que se considere que su nivel de razonamiento debe ser alto y se indica así en la evaluación.

Error Estándar de Estimación. Al aplicar rectas de regresión para estimar el rendimiento en otra variable, como un *Criterio*, si la correlación no es perfecta ($r_{xy} = 1.0$) se comete un error de estimación del criterio. Se utiliza para establecer la posible magnitud del error presente en la estimación (\hat{Y}) y para determinar *Intervalos de Confianza*. Se calcula mediante la fórmula $S_e = S_x \sqrt{1 - r_{xy}^2}$ (en donde S_e es el error estándar de estimación, S_x es la desviación típica de la variable a predecir, y r_{xy}^2 es la correlación entre el predictor y la variable predicha elevada al cuadrado).

Error Estándar de la Diferencia. Estadístico que permite establecer si las diferencias entre los puntajes obtenidos por una persona en dos pruebas o por

dos personas en la misma prueba, son diferentes entre sí en forma lo suficientemente significativa como para afirmar que subyacen diferencias en los conocimientos o habilidades que se están midiendo. Se calcula mediante la fórmula siguiente, en la que S_t es la Desviación Típica de la escala de medición y r_{xx} es el Coeficiente de Confiabilidad:

$$Se_d = S_t \sqrt{2(1 - r_{xx})}$$

Error Estándar de Medida S_{em} . Es un estadístico que estima la posible magnitud del error presente en un puntaje individual. El Error Estándar de Medida indica la diferencia que puede existir entre el Puntaje Obtenido y el Puntaje Real o Verdadero (T), que se supone que refleja la verdadera posición del individuo en un continuo subyacente; esa diferencia se debe a errores de medida. Mientras mayor es el Error Estándar de Medida, menos confiable es la puntuación obtenida. El Error Estándar de Medida es una magnitud tal, que en aproximadamente dos terceras partes de los casos, el puntaje obtenido no diferirá más de un Error Estándar por encima o por debajo del Puntaje Real. En teoría se puede decir, que en 2 de cada 3 casos, el puntaje obtenido se encuentra en una banda que se extiende del Puntaje Real más o menos un Error Estándar de Medida; pero debido a que el Puntaje Real nunca se puede conocer, la práctica establece que, para la interpretación, se puede revertir la relación Puntaje Verdadero – Puntaje Obtenido. Se calcula mediante la fórmula siguiente, en la que S_t es la Desviación Típica de la escala de medición y r_{xx} es el Coeficiente de Confiabilidad:

$$S_{em} = S_t \sqrt{1 - r_{xx}}$$

Escala. La sucesión de números en función de los cuales se realiza la medición. Clásicamente se distinguen cuatro niveles: Nominal, Ordinal, Intervalo, Razón. También se entiende por escalas, las transformaciones de puntajes que se realizan a escalas de propiedades conocidas, tales como *Estaninas Percentiles*, *puntaje Z*, *Puntaje H*, etc.

Estadística Descriptiva e Inferencial. La E. Descriptiva es una rama de la estadística aplicada que permite describir las características de las distribuciones en función de los datos obtenidos. La E. Inferencial es una rama de la estadística que permite realizar inferencias sobre el comportamiento de los estadísticos en la población o un segmento de ella. Ambas son absolutamente necesarias en el trabajo estadístico de los datos psicométricos.

Estanina. Escala de nueve puntos estandarizados. El nombre proviene del inglés Stanine (standard nine), que tiene valores del 1 al 9, con una Media de 5 y una Desviación Típica de 2. Cada Estanina, excepto la 1 y la 9, tiene media desviación típica de ancho, y la Estanina 5 tiene un cuarto de desviación típica por encima y un cuarto por debajo de la media. Cada Estanina contiene un determinado porcentaje de la distribución, de acuerdo a la siguiente Tabla:

	1	2	3	4	5	6	7	8	9
%	4	7	12	17	20	17	12	7	4
% ac.	4	11	23	40	60	77	89	96	100

Evaluación. Es la comparación de una situación actual con un estándar. En el caso de las pruebas psicométricas, los resultados cuantitativos (véase *Medición*), se comparan con estándares para establecer cómo rindió la persona, es decir, evaluar el resultado; la evaluación puede ser la comparación con una *Norma*, o con un *Perfil Psicométrico*. Usualmente la evaluación constituye un juicio de valor.

Factores. En medición psicométrica, factor es un rasgo hipotético que subyace a una variable y que influye sistemáticamente en el rendimiento en los ítems de un test o los puntajes en dos o más tests y por lo tanto, hace que el rendimiento en diferentes tests esté correlacionado. El término "factor" se refiere estrictamente a la variable teórica, que se deriva por un proceso de análisis factorial a partir de una matriz de intercorrelaciones entre tests. Sin embargo, también se utiliza para denotar la interpretación psicológica que se le da a la variable, por ejemplo, el rasgo mental que se asume que está representado por la variable, tales como habilidad verbal, habilidad numérica, etc.

Formas paralelas o alternas. Véase *Pruebas Paralelas*.

Frecuencia. Es el número de observaciones que corresponde a un determinado puntaje en una escala. Usualmente las observaciones son los sujetos que obtuvieron ese puntaje.

Frecuencia Acumulada. Es el número de observaciones acumuladas hasta el límite superior de un determinado puntaje en una escala. Se acumula siempre desde las puntuaciones más bajas en la Escala hacia los más altos.

H (Escala H). Escala a la que se transforman en forma lineal los puntajes con Media de 100 y Desviación Típica 20.

H de Horst. Índice del nivel de *Homogeneidad* de una prueba, que se obtiene mediante la división de la *Varianza* obtenida entre la *varianza* máxima posible, que se da si la *covarianza* entre los ítems es la máxima dados los valores de *P de los Items*.

Homogeneidad. Grado de similitud que existe entre los elementos de una *Población* con relación a ciertos atributos; la *Población* puede ser de *Sujetos*, *Items*, o de *Contenidos* o *Conceptos*. Cuando se trata de sujetos, la *Homogeneidad* se puede establecer en forma comparativa mediante la *Varianza* de una distribución de puntajes obtenidos por dos o más grupos de sujetos; el que tiene la menor *Varianza* es más homogéneo, en este caso hablamos de *Homogeneidad de la Muestra*. Cuando se trata de los ítems de una prueba, se establece mediante *H de Horst*, *KR21*, ó el *Coefficiente Alfa* de Cronbach, que se fundamentan en la *covarianza* entre los ítems; en este caso hablamos de *Homogeneidad de la Prueba*.

Intervalo de Confianza. Recorrido que se establece alrededor de un estadístico (calculado en una *Muestra*) para determinar la ubicación de un parámetro (de la

Población) con un determinado Nivel de Confianza. Se establece en función de los estadísticos de muestreo del estadístico en cuestión, para lo que es usual utilizar las propiedades de la *Distribución Normal*. Por ejemplo, el Intervalo de Confianza de la Media permite estimar la Media de la *Población* (μ), haciendo uso del error estándar de la Media multiplicado por la *z* correspondiente al Nivel de Confianza seleccionado.

Inventario de Personalidad. Es un tipo de Prueba Psicométrica que consiste en un cuestionario o lista de chequeo, usualmente autoadministrada, diseñada para obtener información de variables no intelectuales de un individuo. Los inventarios usualmente evalúan rasgos tales como intereses, actitudes, motivación, etc. (véase Pruebas de Personalidad).

Ítems. Son los estímulos o reactivos que se presentan a los sujetos para que respondan y pongan de manifiesto conductas asociadas con un atributo determinado. Dependiendo del tipo de prueba, los ítems varían sustancialmente. Un tipo común de ítem son los de opciones múltiples, que presentan un problema en el "pie" y varias opciones de respuesta de las cuales sólo una es la correcta (la clave) y las otras son *Distractores*, que sirven para atraer la respuesta de quienes no conocen realmente la respuesta correcta.

Ítems binarios. Son los ítems que aceptan como puntuación solamente 1 ó 0. Por contraste, los *Ítems No Binarios* aceptan puntuaciones de más de 1, o con valores decimales. Los ítems binarios tiene aceptación tanto por su facilidad de corrección, como porque han demostrado gran estabilidad en las puntuaciones y en su nivel de dificultad en muestras sucesivas.

Ítems No binarios. Son ítems que no se puntúan sólo como 1 ó 0. Son usados frecuentemente, pero no exclusivamente, en inventarios de personalidad o en la medición de *Actitudes e Inventarios de Personalidad*, en los que pueden haber escalas de aproximación de cada afirmación a la forma usual de comportamiento del sujeto.

KR21. Estadístico que permite determinar la confiabilidad de consistencia interna de una prueba. Deriva su nombre de las fórmulas propuestas por Kuder y Richardson en *Psychometrika* en los años 30, de las cuales la número 21 fue particularmente exitosa. Tiene estrecha relación con Alfa de Cronchach. Se calcula con la fórmula siguiente en donde KR_{21} es el coeficiente de consistencia interna, m es el número de ítems de la prueba, S_j^2 es la varianza del ítem y S_x^2 es la varianza total del test:

$$KR_{21} = \frac{m}{m-1} \left(1 - \frac{\sum S_i^2}{S_x^2} \right)$$

Longitud del Test y Confiabilidad. Véase *Spearman Brown*.

Media o Media Aritmética. Es el promedio de las puntuaciones obtenidas por un grupo en un test. Se obtiene mediante la suma de las puntuaciones y la división entre el número de ellas.

Media del ítem. Véase *Dificultad de los Items*.

Medición. Consiste en la asignación de numerales a las propiedades de los objetos o eventos de acuerdo a ciertas reglas (Stevens). En medición psicológica, los objetos son las manifestaciones de los atributos psicológicos o constructos. Se fundamenta en varios supuestos: a) la asunción de la existencia de un continuo subyacente con direccionalidad, b) la existencia de un solo puntaje verdadero para cada sujeto, c) la variabilidad entre sujetos con relación al puntaje verdadero y d) la posibilidad de obtener manifestaciones asociadas con la posición que la persona ocupa en el continuo subyacente. De esta manera, se asignan puntuaciones, de acuerdo a ciertas reglas, a las manifestaciones de conducta asociadas al atributo psicológico que se desea medir. Las puntuaciones se interpretan en función de la teoría psicométrica y psicológica, para inferir acerca de la conducta actual, de la futura o de las potencialidades de la persona. La medición usualmente implica la existencia de un propósito y una decisión que debe tomarse con relación a los *Sujetos* de la medición o de los atributos medidos.

Mediana. Es la puntuación en una distribución de puntajes por encima de la cual se encuentra el cincuenta por ciento de las observaciones.

Medidas de Dispersión. Estadísticos que permiten establecer la variabilidad de un conjunto de datos. Por ejemplo: la *Varianza*, y la *Desviación Típica*.

Medidas de Tendencia Central. Estadísticos que permiten conocer la ubicación de la distribución en un continuo, por ejemplo: *Media*, *Mediana* y *Moda*,

Microítem. Véase *Software Psicométrico*

Moda. Es la puntuación que tiene la frecuencia más alta en una distribución. Algunas distribuciones pueden ser bimodales o multimodales.

Muestra. Una parte de una *Población* o universo; un subconjunto de observaciones de una población. Si la población está compuesta de individuos, la Muestra constituye un grupo de individuos provenientes de esa población. Se desea que las Muestras sean representativas de la población para poder inferir los parámetros de la población a partir de los estadísticos de la Muestra. La representatividad se logra asegurando que todos los individuos de la población tienen la misma probabilidad de ser seleccionados para la Muestra, se habla así de muestreo probabilístico. Esto se logra mediante esquemas de muestreos entre los que se incluyen como sus elementos básicos: Muestreo por Azar Simple, Muestreo por Azar Sistemático, Muestreo Estratificado, Muestreo por Conglomerado. Estos esquemas exigen

contar con una descripción exhaustiva de los elementos de la población a fin de identificar cada uno de ellos. Cuando no se cuenta con esa descripción se hace uso de muestreos por cuotas, en los cuales se definen los atributos que influyen

sobre la variable a evaluar y luego se completan "cuotas" de personas en cada una de los niveles de esos atributos.

N. Número de sujetos incluidos en un grupo de sujetos o en un grupo normativo.

Normas. Es la expresión estadística del rendimiento típico de un grupo en una prueba. Usualmente se expresa mediante valores estadísticos o mediante un conjunto de puntajes en una escala de propiedades conocidas. Sirven como referencia para darle significado a los puntajes brutos o puntajes directos, que se obtienen en las Pruebas Psicométricas. Las normas se basan en el rendimiento real de un grupo de personas en una prueba. Dicho grupo se selecciona especialmente para que tenga el mayor grado de similitud, en variables importantes, con los individuos que serán evaluados con la prueba. Debido a que se trata de rendimiento promedio o típico, las Normas no deben ser consideradas como estándares o niveles deseables de rendimiento. También es importante que al construir Normas se definan con claridad las características de los sujetos del grupo, con relación a variables importantes tales como edad, nivel educativo, profesión. Los tipos más comunes de Normas son expresadas en escalas como las Percentiles, Estaninas, Puntaje Z y Puntaje H.

P del ítem. Véase *Dificultad de los Items*.

Pentil. Uno de los niveles en una escala de cinco puntos estandarizados. Tiene valores del 1 al 5 con una media de 3 puntos. La siguiente Tabla muestra las relaciones entre los valores Pentiles, los porcentajes de la distribución y las desviaciones típicas en una curva normal. Nótese que la escala Pentil deja por fuera aproximadamente el 1,2% de los extremos de una distribución normal, es decir, 0,6% de cada lado.

	1	2	3	4	5
%	6	24,2	38,2	24,2	6
% ac	6	30,2	68,4	92,6	98,6
D. T.	-2,5 a-1,5	-1,5 a-0,5	-0,5 a0,5	0,5 a1,5	1,5 a2,5

Percentil. Es un puntaje en la distribución en el cual, y por debajo del cual, se encuentra el porcentaje de casos indicado por el percentil. De esta manera, de un puntaje que coincida con el percentil 85 se dice que iguala o supera al 85% de las personas del grupo normativo y que 15% supera éste puntaje. Es importante destacar, que percentil no tiene ninguna relación con el porcentaje de respuestas correctas dadas a una prueba.

Perfil Psicométrico. Es la representación gráfica de los resultados de un sujeto o de un grupo de sujetos en varios tests. Para ello es necesario que los resultados se expresen en escalas comparables (Percentiles, H, z, Z, etc.). Los perfiles permiten la comparación del rendimiento y así, la identificación de las áreas de fortalezas y debilidades.

Personalidad, Pruebas de. Véase *Pruebas de Personalidad*.

Piso de una Prueba. Es el nivel más bajo que se puede medir con una prueba. Cuando un individuo obtiene una puntaje que coincide o que está muy cerca del puntaje más bajo posible, se dice que la prueba tiene un “piso” muy alto para él y que se le debe administrar una prueba más fácil (véase *Techo de una Prueba*).

Población. Grupo de elementos (que pueden ser personas, ítems, conceptos, etc.) que se definen como pertenecientes a un todo y del cual se obtiene una *Muestra*. También, una colección de todas las observaciones que se pueden identificar de acuerdo a ciertas reglas.

Pruebas Aptitudinales. Son pruebas Psicométricas que miden habilidades y otras características, ya sean innatas o adquiridas, que indican la capacidad de un individuo para aprender o desarrollarse en un área particular académica o laboral, siempre y cuando se le provea entrenamiento adecuado y oportunidades. Las Pruebas Aptitudinales incluyen tanto habilidades académicas generales (llamadas comúnmente habilidad mental o inteligencia), como habilidades especiales como habilidad verbal, numérica, cuantitativa o musical. En general se utilizan para predecir el rendimiento futuro, usualmente en un campo específico, como programas de entrenamiento o en el trabajo.

Pruebas Cognoscitivas. Son Pruebas Psicométricas destinadas a medir aspectos intelectuales de la estructura mental o psicológica del individuo. Incluyen las Pruebas de Conocimientos, las Pruebas Aptitudinales, Las Pruebas de Habilidades y las Pruebas de Personalidad.

Pruebas de Conocimiento. Son pruebas Psicométricas que permiten medir el dominio de información en un campo determinado u área del conocimiento. Por ejemplo, dominio de información sobre técnicas de contabilidad. Son llamadas también Pruebas de Rendimiento.

Pruebas de Diagnóstico. Son tests utilizados para establecer áreas específicas de fortalezas y debilidades, y que se utilizan, en oportunidades, para determinar el origen causal de debilidades y desarrollar programas de intervención individual o grupal. Tales Pruebas permiten establecer medidas de los componentes de habilidades o conocimientos más complejos.

Pruebas de Habilidades. Son Pruebas Psicométricas que miden rasgos cognoscitivos que se han desarrollado a través del entrenamiento, de las oportunidades laborales o de las actividades en la vida cotidiana. Entre ellas se pueden contar las habilidades de lectura, ortografía, redacción. Al igual que en las aptitudes, una definición más amplia y comprensiva de las habilidades probablemente deba incluir también factores asociados con la motivación, los intereses y oportunidades, que permiten que una persona desarrolle sus aptitudes. También es posible percibir las habilidades como la plataforma para el desarrollo en nuevos campos, por lo cual tienen también un carácter de aptitudes.

Pruebas de Personalidad. Son Pruebas Psicométricas destinadas a medir uno o más aspectos no intelectuales de la estructura mental o psicológica del individuo. Permite obtener información acerca de características afectivas del individuo (emocionales, motivacionales, actitudinales, etc.) que se diferencian de sus

habilidades o procesos intelectuales. Las Pruebas de Personalidad incluyen: a) los llamados inventarios de personalidad y de ajuste, que intentan medir la posición de la persona en rasgos tales como dominancia, sociabilidad, introversión, locus de control, autoestima, etc., mediante el uso de respuestas autodescriptivas a una serie de preguntas o afirmaciones. b) Escalas de rating, que exigen que el sujeto se compare, o sea comparado por otros, con respecto a ciertos rasgos. c) Inventarios de opinión o actitudes. Algunos autores también incluyen los inventarios de intereses y creencias como pruebas de personalidad.

Pruebas de Potencia. Son Pruebas diseñadas para medir el nivel de rendimiento independientemente de la velocidad de respuesta. Por ello, usualmente no tienen límite de tiempo o se otorga a los sujetos suficiente tiempo para que puedan terminar sin dificultad. Es usual que los reactivos o ítems se ordenen de manera creciente de acuerdo a su dificultad.

Pruebas de Velocidad. Son Pruebas diseñadas para evaluar habilidades asociadas con la velocidad de respuesta y establecen su varianza en función de la velocidad de respuesta. Usualmente contienen muchos ítems relativamente fáciles, que si se contase con suficiente tiempo serían respondidos por todos los sujetos, pero como deben ser respondidos bajo condiciones de presión de tiempo, usualmente no se logran responder todos.

Pruebas Psicométricas. Véase *Tests Psicométricos*

Pruebas Referidas a Dominio. Son pruebas diseñadas para proveer información sobre conocimientos o habilidades específicas que posee el sujeto de manera directa con relación al dominio que evalúa el test. Tales tests cubren usualmente unidades de contenido y están estrechamente relacionadas con procesos de instrucción. Sus puntajes tienen significado en función de lo que el estudiante sabe o puede hacer, más que en relación con los puntajes obtenidos por un grupo de referencia o grupo normativo. Fueron propuestas inicialmente por Glaser en 1963 y se comenzaron llamando Pruebas Referidas a Criterios, nombre que fue abandonado debido al carácter polisémico en Psicometría del término Criterio, que también puede ser entendido como *Criterio de Validación* y como *Punto de Corte*.

Psicometría. Medición de los atributos psicológicos. En general constituyen métodos experimentales que permiten poner de manifiesto conductas asociadas con los atributos psicológicos; a esas manifestaciones se les asignan numerales que permiten establecer el nivel relativo en que se encuentra ubicada la persona con relación al atributo. Los *Test Psicométricos* son uno de los métodos de la Psicometría.

Puntaje Derivado. Son puntajes que se transforman a escalas de propiedades conocidas. Estas escalas, de las cuales la más conocida con los Puntajes Z, C.I., Desviación H, CEEB y T, tienen una media y una desviación típica determinada lo que permite interpretar las puntuaciones en función de dónde se ubican en la escala. Por ejemplo H tiene una Media de 100 puntos y una Desviación Típica de 20, lo que se representa como H(100,20). Algunos puntajes

Derivados son transformaciones lineales de los Puntajes Directos, mientras que otros son transformaciones No Lineales, como el caso de las Estaninas o T.

Puntaje Directo o Puntaje Bruto. Son los puntajes que se obtienen directamente de la asignación de numerales a las respuestas a una Prueba Psicométrica. Usualmente se obtienen de asignar un punto por cada respuesta correcta. En los casos en que existe una corrección por puntajes al azar, es usual que se calcule mediante la fórmula $X = C - I/(k-1)$ En donde X es el puntaje obtenido, C las preguntas respondidas correctamente, I las respondidas incorrectamente y K el número de opciones de las preguntas de la Prueba Psicométrica.

Punto de Corte o de aprobación. Se refiere al puntaje por encima del cual se considera que el sujeto cumple con ciertos requisitos de ingreso, aprobación de una materia o de selección para un empleo. Pueden ser establecidos fundamentados en estudios que demuestran que a partir de ese punto existe mayor probabilidad de éxito en el trabajo o en los estudios, o de manera más o menos arbitraria, para permitir el ingreso a un cierto número de personas a la institución, como cuando existe un cupo limitado. También ha sido denominado *Criterio*, pero, por el carácter polisémico (muchos referentes o significados) de éste término, es conveniente no utilizarlo en éste contexto.

PsicoMet. Véase *Software Psicométrico*.

Reactivo. Véase *Items*

Recta de Regresión. Transformación lineal de puntajes obtenidos por un grupo de sujetos en dos variables, que se calcula haciendo uso del coeficiente de correlación entre ellas y que permite determinar la puntuación que obtendrá en la otra variable, conociendo la puntuación obtenida en una de ellas. Se calcula haciendo uso de la siguiente ecuación,

$$Y' = r_{xy} \frac{S_y}{S_x} (X - \overline{M_x}) + \overline{M_y}$$

en donde Y' es el estimado de la puntuación en la variable predicha, r_{xy} es el coeficiente de correlación entre ambas, S_x y S_y las desviaciones típicas de ambas variables, X es la puntuación obtenida por un sujeto en la variable predictora, M_x y M_y son las medias de las variables predictora y predicha respectivamente:

Regresión múltiple (véase *Correlación Múltiple*).

Software Psicométrico. Programas de computación utilizados para la *Corrección de Pruebas*, la administración de pruebas por computadora o para el procesamiento de información. En Venezuela se producen y comercializan dos programas: Microitem, que permite el análisis de ítems, la corrección de pruebas de opciones múltiples con ítems binarios y no binarios, y la elaboración de normas, producido por Massimo Di Salvatore de la Universidad Central de Venezuela y PsicoMet, que permite la corrección de pruebas, la elaboración de informes psicométricos grupales e individuales, la administración de una base de

datos de elegibles y la elaboración de normas internas, producido y distribuido por Psico Consult C.A. Ninguno de estos programas permite la administración de pruebas por computadora, pero tienen la ventaja de que permiten el procesamiento de cualquier prueba binaria de opciones múltiples.

Spearman Brown, Fórmula de Profecía. La confiabilidad de un test tiene una relación directa con el número de ítems. Cuando se duplica el número de ítems, la confiabilidad es mayor. Cuando se calcula la confiabilidad por el método de división por mitades, se obtiene la confiabilidad de la mitad de la prueba, por lo que se hace necesario ajustar para la totalidad de los ítems. Se usa así la fórmula de Spearman Brown según la fórmula siguiente, en donde r_{tt} es la confiabilidad obtenida para la mitad de la prueba, r_{tt_n} es la confiabilidad para la prueba total:

$$r_{tt_n} = \frac{2 \cdot r_{tt}}{1 + r_{tt}}$$

Es una forma especial de la fórmula siguiente, en donde n es el número de veces que se aumenta el test.

$$r_{tt_n} = \frac{n r_{tt}}{1 + (n - 1) r_{tt}}$$

Sujeto. Es una de las formas de referirse a las personas que presentan Pruebas Psicométricas o que participan en un experimento.

Supuestos de la medición (véase *Medición*).

T. Transformación no lineal derivada de z_n con Media 50 y Desviación Típica 10.

Tabla de Especificaciones. Son tablas que permiten definir el contenido de una prueba; son particularmente útiles para la elaboración de *Pruebas de Conocimiento* y para determinar su validez. Usualmente se establecen como tablas de doble entrada (aunque pueden incluir más de dos dimensiones), una de las cuales indica el contenido del proceso.

de aprendizaje y la otra alguna definición de las habilidades que la persona logró durante ese proceso (es usual usar la Taxonomía del Dominio Cognoscitivo de B. Bloom y otros).

Tablas de Expectación. Son tablas que muestran la relación entre los puntajes de un test predictivo y el rendimiento en un criterio determinado. La relación se puede expresar de maneras diferentes tales como: a) El promedio de rendimiento en el criterio. b) El porcentaje de casos en cada uno de los niveles sucesivos. c) La probabilidad de obtener una puntuación en el criterio, dada una puntuación en la Prueba. Las Tablas de Expectación son comunes para la predicción del éxito académico o laboral, son de fácil elaboración y comprensión. Se diferencian las Tablas de Expectación Individual de las Tablas de Expectación Institucional, ya que en las primeras se indica la probabilidad de obtención de un determinado resultado en el *Criterio*, mientras que en las segundas se plantea

solamente la probabilidad de éxito o fracaso en el mismo. El nivel más avanzado incluye varias variables predictoras y se construyen a partir de rectas de regresión múltiple.

Techo de una Prueba. Es el nivel más alto de habilidad que se puede medir con un test. Cuando un individuo logra un puntaje que coincide o que está cerca del puntaje más alto posible, se dice que el test tiene un "techo", muy bajo para él y es necesario administrarle una prueba que sea más difícil (véase *Piso de una Prueba*).

Tendencia Central (medidas de). Las medidas de tendencia central proveen una representación del rendimiento de un grupo, indica la tendencia del grupo de medidas. Las más conocidas son la media, la mediana y el modo.

Tests Psicométricos. "Un test psicológico constituye esencialmente una medida objetiva y tipificada de una muestra de conducta" Anastasi, Anne. Rodríguez y Feliú lo definen como "... un conjunto de estímulos seleccionados haciendo uso de una fundamentación teórica, psicológica y psicométrica, en función de un propósito. Estos estímulos son presentados a sujetos en condiciones controladas con el objeto de poner de manifiesto conductas relevantes que debidamente interpretadas con base a normas y/o criterios permiten la toma de decisiones respecto a un individuo, un grupo o un proceso."

Transcripción de Resultados. Es la acción de registrar en un programa de computación (por ejemplo, el PsicoMet) las respuestas dadas por los sujetos a las preguntas de una Prueba Psicométrica.

Transformación de puntajes. Los *Puntajes Directos* que se obtienen por la asignación de numerales a las respuestas de los sujetos, usualmente no tienen una significación inmediata. Por ello se realizan transformaciones a escalas de propiedades conocidas, que permiten contar con un resultado más fácilmente interpretable en función del rendimiento de grupos normativos. Se reconocen al menos dos tipos de transformaciones: lineales y no lineales. Las lineales se realizan mediante ecuaciones que utilizan estadísticos como la media y la desviación típica y no modifican la forma de la distribución (por ello existe una relación rectilínea entre los puntajes directos y sus puntajes transformados); las más conocidas son z , Z , y H . Las no lineales hacen uso de ciertas propiedades, como el número de casos acumulados en la distribución y el puntaje transformado (por ello existe una relación no lineal entre los puntajes directos y sus puntajes transformados); las más conocidas son los *Percentiles*, *Estaninas*, *Pentiles*, z_n y T .

Validación Cruzada. Procedimiento mediante el cual la validez de una prueba establecida en una muestra poblacional se verifica con otra muestra proveniente de la misma población. Es usual derivar rectas de regresión cuando se hacen estudios de validez; al realizar un estudio de Validación Cruzada, se utiliza esa recta para determinar los mejores valores predichos en otra muestra de la población, y luego se verifica qué tan cerca está la predicción de los resultados obtenidos en el *Criterio de Validación* haciendo uso de métodos correlacionales.

Validez. La medida en que el test mide lo que se pretende que mida, es decir, que sirva para lo que se pretende utilizar. El término validez, por lo tanto, tiene

connotaciones diferentes dependiendo el uso para el cual se propone y para diferentes tipos de tests. Se habla así de *Validez Aparente*, *Validez de Constructo*, *Validez de Contenido*, *Validez Empírica o Criterial*,

Validez Aparente. Es la medida en que el test aparenta ser un instrumento eficiente de medición. En realidad tiene más que ver con lo que los *Sujetos* piensan del instrumento y la situación en que se aplica y la credibilidad que pueden tener en el proceso de evaluación y la atención que prestarán a la misma.

Validez de Constructo. Pretende establecer hasta qué grado la varianza total observada se puede explicar en función de variables (*Constructos*) derivados de una teoría. Para realizar estudios de validez de constructo, es necesario contar con una buena descripción teórica, que permita derivar hipótesis y probarlas haciendo uso del instrumento. En la medida en que se obtienen los resultados predichos por la teoría, se dice que hay validez de constructo.

Validez de Contenido. Determinación de si una prueba mide lo que se supone que debe medir con relación a la Tabla de Especificaciones que describe el universo de evaluación (tópicos de contenido, habilidades desarrolladas, etc.). Se establece cuando un grupo de expertos analiza cualitativamente los ítems y determina dónde se ubican con relación a las celdas de la Tabla de Especificaciones. Si existe coincidencia entre los juicios de los expertos y lo que se prevé que mida cada ítem, se dice que hay validez de contenido. Es decir, se trata de determinar qué grado de representatividad tiene la prueba con respecto a universo de conceptos que se pretende medir. Se aplica de manera particular en las *Pruebas de Rendimiento*, que evalúan un contenido determinado.

Validez Empírica o Criterial. Establece la relación entre el rendimiento en el test y las puntuaciones obtenidas en una variable externa (criterio) que se supone mide el mismo atributo que el test. Hace uso de métodos correlacionales (véase *Correlación*), que permiten establecer el grado de asociación entre ambas variables. La validación empírica se fundamenta en estudios que tienden a ser de carácter aplicado, puesto que a menudo interesa solamente la capacidad predictiva de los resultados de la prueba con relación al criterio. Cuando el criterio se toma algún tiempo después de la aplicación de la prueba, se habla de validez predictiva, como cuando se trata de predecir el rendimiento universitario a partir de una prueba de aptitud académica. Cuando se toman la medida en el test y en el criterio en forma simultánea se habla de validez concurrente, como cuando se trata de realizar un diagnóstico de los conocimientos de un grupo de personas.

Validez predictiva y concurrente. Véase *Validez Empírica o Criterial*.

Varianza. Medida de la dispersión de los puntajes en una distribución. Se calcula mediante el promedio de los desvíos de cada puntaje con respecto a la *Media* elevados al cuadrado.

Varianza del ítem. En *ítems binarios* se determina mediante la multiplicación de pxq , en donde p es la proporción de respuestas correctas y q la de respuestas incorrectas al ítem. En este tipo de ítems, su valor está contenido entre 0 y 0,25. En ítems no binarios, tanto para su cálculo como para establecer sus límites, es necesario conocer el recorrido de la escala de los ítems. La raíz cuadrada de la varianza del ítem constituye la desviación típica del ítem.

Velocidad, Pruebas de. Véase *Pruebas de Velocidad*.

z o puntaje Estándar. Es un término general que se refiere a la transformación de puntajes a escalas de propiedades conocidas. En el caso de los puntajes z (minúscula) es una transformación referida al número de desviaciones con respecto a la media a la que se encuentra un determinado puntaje. Se calcula mediante la fórmula siguiente (en donde X es un puntaje determinado, \bar{X} es la Media y D.T. es la Desviación Típica) $z = (X - \bar{X}) / D.T.$

Z (Escala Z). Transformación lineal de Media de 50 puntos y Desviación Típica de 10 puntos.

z_n (z normalizada). Transformación a una escala de Media = 0 y Desviación Típica = 10, fundamentada en las proporciones acumuladas hasta el punto medio de la clase haciendo uso de la Tabla z de áreas bajo la curva normal. Es una transformación no lineal.

Algunas Fuentes bibliográficas consultadas para la elaboración de este Glosario

Anastasi, Anne. Urbina, Susana. (1998) Tests Psicológicos. México, Prentice Hall

Guilford, J.P. (1954). Psychometric Methods New York, McGraw Hill.

Guilford, J.P. (1960) Fundamental Statistics in Psychology and Education. New York, McGraw Hill.

Kerlinger, Fred (1973) Foundations of Behavioral Research. N.Y. Holt, Rinehart and Winston.

Kirk, Roger. (1969). Experimental Design: Procedures for the Behavioral Sciences. Belmont, California, Brooks Publishing Co.

Magnusson, David (1976) Teoría de los Tests. México, Editorial Trillas.