



UAEM

Universidad Autónoma  
del Estado de México



*Centro Universitario UAEM Ecatepec*

*MONOGRAFÍA PARA EL DISEÑO, ELABORACIÓN Y  
ANÁLISIS DE LOS REACTIVOS DE UN INSTRUMENTO.*

UNIDAD DE APRENDIZAJE:  
ELABORACIÓN DE INSTRUMENTOS  
Programa por Competencias

Licenciatura en Psicología

Dr. en Ed. Carlos Saúl Juárez Lugo

Agosto de 2015

## INDICE

Datos curriculares de la unidad de aprendizaje	3
Mapa Curricular	4
Mapa conceptual	5
Propósito de Aprendizaje	6
Estrategia General	6
Presentación	8
Introducción	11
La elaboración de un instrumento piloto: diseño, elaboración y análisis de los reactivos	13
I.    Concepto	14
II.   Dimensiones	15
III.  Indicadores	15
IV.  Índices	15
V.   Redacción del reactivo	18
VI.  La escala de medición del reactivo	19
Metodología para determinar la calidad de los reactivos de un instrumento	21
I.    Análisis del examen aplicado	22
II.   Elaboración de la hoja de cálculo	22
III.  Definición de la escala de clasificación de la dificultad del ítem	25
IV.  Análisis de reactivos:	
índice de dificultad	26
índice de discriminación.	33
V.   Análisis de la confiabilidad de las preguntas de un instrumento: alfa de Cronbach.	41
VI.  Análisis de los reactivos del EXANI-I	46
Bibliografía	48

# MONOGRAFÍA PARA EL DISEÑO, ELABORACIÓN Y ANÁLISIS DE LOS REACTIVOS DE UN INSTRUMENTO.

## DE LA UNIDAD DE APRENDIZAJE: ELABORACIÓN DE INSTRUMENTOS

### Licenciatura en Psicología

#### UNIDAD DE APRENDIZAJE: ELABORACIÓN DE INSTRUMENTOS

Academia: Estadística y Construcción de Instrumentos

Núcleo de formación: Metodológica instrumental

Competencia: Habilidad - investigación

Modalidad: Presencial

Créditos: 8

Horas teóricas: 2 horas s/m

Horas prácticas: 2 horas s/m

Duración total del curso: 64 horas

Se imparte en el cuarto periodo de la licenciatura.

Unidades de Aprendizaje antecedentes: ninguna.

Unidad de Aprendizaje subsecuente: Taller de elaboración de instrumentos.

Conocimientos y habilidades previas: Estadística, estadística aplicada, investigación cuantitativa, entrevista y taller de la entrevista.

Unidades de aprendizaje simultáneas:

Las unidades académicas por las que optaron los alumnos en función de su ruta académica.

Ubicación de la Unidad de aprendizaje “ELABORACIÓN DE INSTRUMENTOS”

MAPA CURRICULAR DE LA LICENCIATURA EN PSICOLOGÍA

1 periodo	2 periodo	3 periodo	4 periodo	5 periodo	6 periodo	7 periodo	8 periodo	9 periodo	10 periodo
Socialización y Contexto		Fundamentos de Psicología Clínica	Estancia Integrativa Básica	Técnicas e instrumentos psicológicos	Estrategias de Aprendizaje	Psicología del Trabajo	Psicología del Trabajo	Evaluación Profesional I	Evaluación Profesional II
Psicofisiología I	Psicofisiología II	Procesos Psicológicos en el Desarrollo II	Psicología Social	Psicopatología del Niño	Integración de Recursos Humanos	Introducción a la Psicoterapia	Introducción a la Psicoterapia	Estancia Integrativa Profesional I	Estancia Integrativa Profesional II
Procesos Psicológicos Básicos	Procesos Psicológicos Superiores I	Procesos Psicológicos Superiores II	Elaboración de Instrumentos	Taller de Elaboración de Instrumentos	Proceso Grupal	Psicología Comunitaria	Psicología Comunitaria	Optativa	Optativa
Teorías de la Personalidad	Procesos Psicológicos en el Desarrollo I	Fundamentos de Psicología Educativa	Fundamentos de Psicología Organizacional	Taller de la entrevista psicológica	Psicopatología del Adulto	Estancia Integrativa Metodológica	Estancia Integrativa Metodológica	Optativa	Optativa
Derechos Humanos	Metodología de la Ciencia	Investigación Cuantitativa	Investigación Cualitativa	Administración	Optativa	Educación Especial	Educación Especial	Optativa	
Epistemología	Estadística	Estadística Aplicada	Optativa	Orientación Educativa	Optativa	Optativa	Optativa		
		Entrevista	Optativa	Optativa	Optativa	Optativa	Optativa		
		Optativa Optativa		Optativa Inglés C1	Inglés C2	Optativa	Optativa		

# MAPA CONCEPTUAL DE LA UNIDAD DE APRENDIZAJE “ELABORACIÓN DE INSTRUMENTOS”

## **Temas**

1. Fundamentos teóricos en la medición de instrumentos de medición para planear uno propio.
2. Bases metodológicas para construir un instrumento y la generación de reactivos.

## **Objetivos**

1. Conocerá los fundamentos teóricos en la medición en Ciencias Sociales para planear la construcción de un instrumento propio.
2. Redactar de manera clara los reactivos que constituirán el instrumento, teniendo como base la aplicación de redes semánticas naturales y seleccionarán las opciones de respuesta más adecuadas para el constructo a medir.
3. Someter el instrumento inicial a la consulta de expertos y realizará las modificaciones pertinentes para pilotearlo.

## **PROPÓSITO DE APRENDIZAJE**

El objetivo principal de la unidad de aprendizaje de Elaboración de Instrumentos, es que los estudiantes conozcan las bases teóricas para planear, diseñar y elaborar un instrumento válido y confiable, como por ejemplo una encuesta de opinión o una escala de actitud, identificando los elementos siguientes: planeación de un instrumento, obtención de reactivos, construcción del instrumento, consulta a expertos y pilotaje.

## **ESTRATEGIA GENERAL**

Debido a la naturaleza del contenido de la unidad de aprendizaje que involucra directamente al quehacer del alumno en cuanto a sus acciones específicas para enfrentar el aprendizaje, se sugiere centrar las acciones educativas en la propia experiencia de los participantes y el análisis del contexto educativo en el que se encuentran insertos. La primera acción recomendada en la sesión inicial, es el encuadre que comprende al menos la presentación del instructor y de los participantes. Después es necesario explorar las expectativas que tienen los alumnos con respecto al contenido de esta unidad de aprendizaje (U.A), ubicándola en el mapa curricular de la Licenciatura en Psicología y destacando su importancia en la formación profesional. Inmediatamente es conveniente recuperar los conocimientos previos con los que cuentan los alumnos, resultado de la experiencia acumulada en su larga trayectoria académica en distintos niveles educativos: conocimiento de algunas cédulas para realizar encuestas, participación en encuestas, por ejemplo censos poblacionales, estudios de mercado, encuestas de servicios y de opinión, si ha respondido pruebas psicológicas, etc. Consultar si algún alumno ha sido encuestador. Explorar

conocimientos previos relativos a estadística descriptiva y aplicada, investigación cuantitativa y de psicología en general.

Como estrategia de la instrumentación didáctica desde un enfoque constructivista y de las competencias académicas, se sugiere orientar el proceso de enseñanza- aprendizaje con base en un diagnóstico, planeación, realización y evaluación continua, además de establecer estrategias de lectura comentada, análisis de instrumentos de recolección de datos, pruebas educativas y pruebas psicológicas, trabajo en equipos colaborativos, la aplicación, análisis y evaluación de instrumentos adecuados al contenido.

Se recomienda que el alumno inicie su colección de cédulas de instrumentos de cualquier tipo y naturaleza, con la finalidad de contar con ejemplos que le permitan analizar criticar las preguntas que conforman las cédulas y sugerir acciones para mejorarlas. Los ejemplos de instrumentos pueden ser aquellos que se encuentra en revistas para jóvenes, las encuestas de servicio en restaurantes o cines hasta las pruebas psicológicas que se encuentran su biblioteca.

## PRESENTACIÓN

Las propuestas curriculares desarrolladas en el último siglo han sufrido un proceso de transformación de una visión centrada en contenidos temáticos hacia una visión centrada en el alumno. En la actualidad ya no es suficiente adquirir ciertos conocimientos o dominar unas técnicas, aunque sean funcionales en la práctica, sino que es necesario articular al menos tres componentes: el actitudinal, el procedimental y el conceptual, para que el actuar del alumno, ante una situación concreta, sea eficiente y adecuada, es decir que el alumno sea competente.

Las competencias se definen como actuaciones integrales ante actividades y problemas del contexto, con idoneidad y compromiso ético, integrando el saber ser, el saber hacer y el saber conocer en una perspectiva de mejora continua. Están constituidas por procesos cognitivo-afectivos subyacentes y por procesos públicos y observables en tanto implica una acción de sí para los demás y/o el contexto. En las competencias está implícita la capacidad de movilizar varios recursos de naturaleza cognitiva como son los conocimientos, las habilidades y actitudes, para hacer frente a un tipo de situaciones. Estos recursos son los que la competencia moviliza, integra y organiza en una situación única, con la posibilidad de ser tratada como análoga con otras ya conocidas; precisamente es este continuo ejercicio de la competencia lo que la forma, al pasar por operaciones mentales complejas, sostenidas por esquemas de pensamiento que permiten determinar y realizar la acción idónea en una constante supervisión así como la evaluación de su actuación.

Es aquí donde destaca la importancia del estudio por parte del psicólogo en formación, de los contenidos de la presente Unidad de Aprendizaje **Elaboración de Instrumentos y en particular de la elaboración y el análisis de los reactivos de un instrumento de recolección de datos**. El empleo de instrumentos de medición para la recolección de datos



va más allá de la sola selección y aplicación de una prueba. Es necesario que el estudiante de psicología conozca de manera puntal el proceso de elaboración de un instrumento que cumpla con los criterios psicométricos establecidos por la comunidad científica, la delimitación de lo que se va a investigar, la estructura de las dimensiones, índices, indicadores e ideas para elaborar los ítems correspondientes así como la redacción pertinente del enunciado que formará el reactivo que corresponda a un tipo de pregunta particular.

Además el psicólogo en formación tiene que conocer los criterios a seguir para dar forma al diseño de la cédula, misma que como instrumento piloto, será aplicado a una muestra representativa de la muestra real objeto de estudio. La aplicación debe cumplir con un conjunto de requisitos que den certeza a los datos recolectados, de acuerdo a criterios metodológicos relativos a la muestra y población, para después pasar a la etapa de captura y análisis estadístico de la información en donde debe emplear e interpretar procedimientos estadísticos como los de confiabilidad y validez, índices de discriminación y dificultad del ítem, así como de consistencia interna del propio instrumento.

Por último el estudiante, de acuerdo a los resultados y su interpretación que responden a un proceso de indagación, tiene que tomar una decisión acerca de la pertinencia del instrumento elaborado como un recurso viable para medir un constructo psicológico en un conjunto de sujetos. Esta decisión fundamentada, tiene que ser reportada en un texto científico para dar certeza de su aporte a la comunidad psicológica.

La Unidad de Aprendizaje “**Elaboración de Instrumentos**” es un saber indispensable en la formación de profesionales altamente calificados en el campo de la psicología educativa. Según el perfil de egreso se considera prioritario que los psicólogos conozcan y

manejen de forma eficiente los elementos teóricos, procedimentales, contextuales y actitudinales relacionados con la psique del individuo.

Precisamente el marco referencial del presente material didáctico responde al modelo curricular basado en competencias, el cual surge como una propuesta para vincular la educación superior con el campo laboral, y de forma paralela fomentando el desarrollo de habilidades cognitivas superiores. Al planearse como objetivo una educación integral, centrada en el aprender a aprender, el plan de estudios de la Licenciatura en Psicología sugiere establecer ambientes de aprendizaje que permitan procesar los conocimientos, las habilidades, las actitudes y los valores necesarios para un desempeño académico satisfactorio construyendo así saberes útiles al individuo y la sociedad. En este sentido, el contenido del presente material ofrece una alternativa para abordar el tema de análisis de reactivos, utilizando métodos y estrategias que facilitan que los actores educativos (el docente y los alumnos) conozcan, comprendan, analicen y construyan críticas y propuestas en torno a los contenidos abordados en la U.A.

Por lo que de acuerdo al programa de la unidad de aprendizaje se pretende que el alumno, obtenga tres competencias específicas: **competencia cognitiva**: conocer los conceptos básicos para el análisis de reactivos como parte de la elaboración de los instrumentos en Ciencias Sociales, **competencia de habilidad**: elaborar un instrumento válido y confiable y **competencia de actitud**: aplicación piloto del instrumento, para analizar el trabajo realizado a la luz de una pequeña muestra de sujetos con las características similares a las que tendrá su muestra para el Taller de la Unidad de aprendizaje consecuente.

## INTRODUCCIÓN

Antes de iniciar la explicación del proceso del análisis de reactivos que conforman una prueba, es importante señalar que esta etapa forma parte de un proceso más amplio que inicia con la delimitación del área de conocimiento de la cual se desea conocer (medir) qué tanto sabe un individuo de ese conjunto de saberes y quiénes, del conjunto de individuos, debería obtener un puntaje alto en la prueba y por el contrario quiénes de ellos obtendrían un bajo puntaje.

Una vez delimitada el área de conocimiento que se desea medir se procede a identificar los elementos que componen ese saber. Enseguida se identifican los indicadores y los índices, para finalmente iniciar con la redacción pertinente del conjunto de ítems o reactivos que en su conjunto formaran el instrumento de medición.

El instrumento resultado de esta parte del proceso, se aplica a una muestra de individuos que son representativos de una población, se obtiene los datos, se procede a su captura en una base o matriz de datos para su posterior análisis, tema del presente material didáctico. Para finalmente obtener una prueba válida y confiable que mida el conjunto de saberes que se desea medir.

Esta breve descripción es la que, por ejemplo, el Centro Nacional para la Evaluación de la Educación Superior (CENEVAL) emplea para elaborar las pruebas conocidas como EXANI, misma que tomaremos como ejemplo para demostrar la forma en que se realiza el análisis estadístico de los ítems de un conjunto de asignaturas que a su vez, contribuye a la construcción de un instrumento más amplio. Para ello, al lector le serán proporcionadas con este documento un conjunto de diez bases de datos que corresponden a la aplicación de los módulos que el EXANI-I incluye para el ingreso al nivel medio superior, a saber: *Biología, Español, Formación Cívica y Ética, Física. Geografía, Historia, Matemáticas, Química,*

*Razonamiento Lógico Matemático y Razonamiento Verbal.* Además se incluye una base de datos que corresponde a la versión del EXANI-I que un grupo de estudiantes de psicología elaboró como práctica en la unidad de aprendizaje Elaboración de Instrumentos.

Para demostrar el procedimiento básico del análisis de reactivos de un instrumento trabajaremos con el archivo que corresponde al área de español. Una vez demostrado el procedimiento, el alumno puede ejercitar su conocimiento con las demás áreas para finalmente realizar el análisis de reactivos con la base EXANI-I 2013.

LA ELABORACIÓN DEL INSTRUMENTO PILOTO:  
DISEÑO, ELABORACIÓN Y ANÁLISIS DE LOS REACTIVOS.

- I. Delimitación del concepto
- II. Establecimiento de las dimensiones del concepto
- III. Identificación de los indicadores
- IV. Elaboración de los índices
- V. Redacción del reactivo
- VI. La escala de medición del reactivo

## I. Concepto

La delimitación del concepto o “idea inicial” se obtiene de la revisión de la literatura existente, de la teoría y de la práctica que se ha publicado respecto al tema del cual se pretende elaborar un instrumento (Aiken, 2003). La revisión de la literatura especializada es de suma importancia. Nos indica qué se ha estudiado del tema y cómo se ha estudiado, a través de qué instrumentos y con qué variables se ha vinculado el constructo principal (Thorndike y Hagen, 2003). Al identificar la estructura del constructo o concepto de acuerdo a su marco teórico empírico, el especialista que elabora un instrumento tendrá la posibilidad de identificar las dimensiones de las cuales se compone éste (Adkins, 2008).

Un ejemplo de este proceso se muestra en artículo de Gargallo, Suárez, y Pérez (2009) titulado “El cuestionario CEVEAPEAU. Un instrumento para la evaluación de las estrategias de aprendizaje de los estudiantes universitarios”.

Al revisar la literatura es indispensable poner atención en los instrumentos que utilizaron los investigadores e identificar (Tornimbeni, Pérez y Olaz, 2014):

- El propósito del instrumento y analizar si es similar al nuestro,
- Las variables que mide,
- Los dominios de contenido de las variables (dimensiones, indicadores, índices),
- A qué individuos fue administrado,
- Condiciones de administración del instrumentos (lugar, contexto, tiempo, etc.),
- Información sobre confiabilidad y validez.

Gargallo et. al., (2009) muestra como desarrolló su instrumento, las críticas a otros que ya existen y que miden el mismo constructo. Así también indica la estructura atendiendo al concepto índices, indicadores y enunciados de acuerdo al fundamento teórico.

## II. Dimensiones

Entendamos a las dimensiones como aquellos elementos relevantes y mutuamente excluyentes que componen el concepto (Esteban Ruiz y Cerezo, 1996). Por ejemplo si trabajamos el constructo inteligencia desde la perspectiva de las inteligencias múltiples, las dimensiones serían:

- Inteligencia lingüística,
- Inteligencia musical
- Inteligencia lógico-matemática
- Inteligencia kinestésica-corporal
- Inteligencia interpersonal e intrapersonal.

Podemos observar que “cada una de las inteligencias” puede ser “observada” como diferente a las otras pero pertenecen a un constructo más amplio.

## III. Indicadores

Los indicadores se entienden como las aquellas características que son propias o compartidas por cada una de las dimensiones establecidas (Macia, 2010; Yela, 1996a). En el ejemplo mencionado si hablamos que vamos a medir el constructo inteligencia, de la dimensión *lógico-matemática*, ésta se caracteriza por la presencia (o ausencia) de habilidades en la solución de problemas y en el pensamiento científico, por el manejo de símbolos o fórmulas, organización gráfica de la información, secuencias numéricas y cálculos.

#### IV. Índices

Los índices se consideran como la idea o enunciado a partir del cual se van a generar las preguntas concretas que se les hacen a los sujetos (Baessa y Arroyave, 1996). Por ejemplo la idea a partir de la cual se va a formular una pregunta para explorar la inteligencia lógico-matemática sería la de *resolver una suma con números decimales*. De esta forma podemos elaborar una pregunta *ad hoc* al enfoque teórico de nuestro constructo, por ejemplo:

a)	$\begin{array}{r} 3.50 \\ + \\ \underline{2.70} \end{array}$
b)	$\begin{array}{r} 3.50 + 2.70 = \\ 3.51 \end{array}$
c)	Si compras un lápiz con un valor de \$2.70 y una pluma con un costo de \$3.50 ¿Cuánto tienes que pagar en total?



En concreto el proceso se muestra en la tabla 1.

<i>CONCEPTO</i>	“Idea inicial” bibliografía existente teoría y práctica	INTELIGENCIA Capacidad para resolver problemas (Gardner).
<i>DIMENSIONES</i>	Elementos que componen el concepto.	Inteligencia: Verbal Espacial Manual Numérica Abstracta
<i>INDICADORES</i>	Características que son propias o compartidas por la dimensión.	Inteligencia lógico -matemática. (Habilidades en la solución de problemas y en el pensamiento científico). Símbolos o fórmulas, organización gráfica, secuencias numéricas, cálculos.
<i>INDICES</i>	Preguntas concretas que se les hacen a los sujetos	Resolver una suma con números decimales.
<i>FORMA EN QUE SE PREGUNTA</i>	3.70 + <u>2.70</u>	3.50 + 2.70= _____
<i>PREGUNTA TENTATIVA</i>	Si compras un lápiz con un valor de \$2.70 y una pluma con un costo de \$3.50 ¿Cuánto tienes que pagar en total?	

Tabla 1. Etapas para la elaboración de los reactivos de un instrumento.

## V. Redacción del reactivo

El siguiente paso es la redacción de una serie de enunciados que expresen con claridad la característica que se quiere explorar. El número de enunciados, de preferencia, debe ser tal cantidad que permita la selección del mejor ítem que exprese con claridad y sencillez la característica que deseamos medir (Kerlinger y Lee, 2001). El número total de enunciados para la primera versión del instrumento que va a presentarse a la muestra inicial de *jueces*, varía conforme al autor. Una cantidad superior de 3 enunciados por 1 idea, permite una mayor rigurosidad en los procesos estadísticos para determinar el poder discriminativo de cada ítem (Tornimbeni et al., 2014). Recomendamos elaborar al menos tres reactivos para cada índice.

Una vez que tenemos los enunciados procedemos a revisar tanto su forma como su contenido. En este último caso el investigador debe acudir a especialistas en el tema que pretendemos medir para que juzguen sí el contenido del enunciado es congruente al objeto que deseamos medir (Adkins, 2008). Este paso se realiza una vez que estamos seguros de que los errores en la forma fueron corregidos.

Para verificar la forma existen varias recomendaciones que enseguida se citan (Briones, 2003; Padua, 1996, Aiken, 2003).

El enunciado debe ser:

1. Univoco, cada una de las cuestiones que la integran tiene tan sólo una respuesta correcta, inconfundible y precisa.
2. Inequívoco, su lenguaje es claro y preciso, evitando interpretaciones falsas.
3. Adaptado, su contenido está acorde con el objeto de estudio y participante.
4. Suficiente, al contener todos aquellos aspectos que se consideran como fundamentales.
5. Ser formulado en un lenguaje sencillo, claro y directo.
6. Reflejar un juicio favorable o desfavorable con respecto a la actitud.
7. Ser pertinente a la variable en estudio. La claridad de la definición del concepto que se mide, facilita redactar enunciados congruentes.

8. Tener una sola interpretación posible para los entrevistados.
9. Vigilar que la construcción gramatical de la frase no sea mayor a 20 palabras.
10. Usar palabras que el entrevistado comprenda.
11. Evitar negaciones (la palabra no) y dobles negaciones.
12. Evitar adjetivos y adverbios no específicos (muchos, en ocasiones) o universales (todos, siempre, ninguno o nunca).

## **VI. La escala de medición del reactivo**

Un elemento esencial en la redacción del reactivo es la elaboración de la alternativa de respuesta esperada. Esta puede adquirir la forma dicotómica o nominal de si – no, acierto – error, cierto – falso (presencia o ausencia del rasgo) como ocurre en un examen de conocimiento. También puede adquirir la forma de una escala como en los instrumentos que miden la actitud (Magnusson, 2009). La redacción de cada uno de los reactivos debe coincidir con la expresión de la escala de la cual, la persona va a escoger sólo una alternativa para indicar su respuesta y de la cual depende directamente el tipo de análisis estadístico.

Veamos las características de cada uno de los tipos de nivel de medición su relación con el análisis estadístico y un ejemplo (Kerlinger y Lee, 2001; Magnusson, 2009; Nunnally, 1991).

Las características del nivel de medición **nominal** son:

- Tiene dos o más categorías del ítem o variables
- Las categorías no tienen orden jerárquico.
- Indica diferencias respecto a una o más características.

- Los elementos que se incluyen en cada categoría se cuentan, no se miden.
- Los números utilizados tienen una función de clasificación.
- No se pueden manipular aritméticamente.

El análisis estadístico asociado a este nivel de medición es: la moda, la frecuencia y el coeficiente de contingencia.

Un ejemplo sería la asignación de números a la variable género (Hombre – Mujer).

Hombre = 1

Mujer = 2

Cuando se afirma que en el nivel de medición nominal los valores no los podemos manipular aritméticamente se refiere a que si tenemos dos hombres y los sumamos ( $1 + 1 =$  ¿?) el resultado es dos, pero este número no significa que al “sumar dos hombres” el resultado real sea “una mujer”. La operación aritmética es correcta pero la representación, interpretación y manipulación de los valores numéricos no lo es.

Las características de nivel de medición **ordinal** son:

- Hay varias categorías de ítem o variable.
- Mantienen un orden de mayor a menor.
- Las categorías indican jerarquía, orden no magnitudes.
- Los elementos que se incluyen en cada categoría se cuentan, no se miden.
- No se conoce con precisión qué tanto es la diferencia entre una y otra.
- No se pueden realizar operaciones aritméticas.

Como ejemplo se puede citar el ordenar un conjunto de juguetes del más chico al más grande. Al juguete más pequeño se le asigna 1, al siguiente juguete de mayor tamaño se

le asigna 2, el que le sigue en tamaño 3, así sucesivamente. La escala de medición ordinal indica un orden pero no la magnitud de la diferencia de ese orden. Sabemos que el juguete 2 es mayor que el juguete 1 y el juguete clasificado como 4 es mayor que los juguetes 3, 2, 1. Pero la magnitud de esa diferencia no la podemos conocer con exactitud. Además tampoco podemos sumar o restar el valor número de los juguetes. La estadística asociada a este nivel de medición es la media, el percentil, Spermán V, Kendall T, Kendall W.

Las características del nivel de medición conocido como **intervalos** son las siguientes:

- Hay orden o jerarquía entre las categorías.
- La distancia entre cada intervalo es igual y constante en la medición.
- Existe una unidad de medida, que se utiliza para “medir” y no solamente contar.
- El cero es arbitrario, no es real.
- Realiza operaciones aritméticas y sus derivaciones.

Podemos citar como ejemplo las respuestas a una prueba de 30 ejercicios matemáticos aplicado a estudiantes universitarios. Podemos afirmar que la distancia (intervalo) entre el alumno que obtuvo 15 afirmaciones correctas y uno que obtuvo 16, es la misma distancia (intervalo) que entre el alumno que obtuvo 20 afirmaciones correctas y el que obtuvo 21, en ambas comparaciones el intervalo es 1. Podemos realizar operaciones aritméticas con las respuestas correctas e incorrectas, inclusive podemos obtener un promedio del número de aciertos / errores del total de participantes. Sin embargo el que un participante obtuviera un puntaje próximo a cero o cero en la aplicación de la prueba, no significa que “no sabe nada de aritmética”, es decir el cero no es real, es un punto de referencia. Otros ejemplos son el

empleo de un termómetro o una prueba de inteligencia, en todos los casos mencionados podemos utilizar estadística descriptiva con la media, la desviación estándar y la **varianza**, estadística inferencia como t o anova al comparar el desempeños entre el sexo de los participantes o de varias escuelas.

Por último tenemos el nivel de medición de **razón**, sus características son:

- La distancia entre cada intervalo es igual iguales entre las categorías.
- Se pueden realizar operaciones aritméticas y sus derivaciones.
- El cero es real, es absoluto. Indica que no existe el atributo

Como ejemplo tenemos el tiempo de exposición a la televisión de un grupo de alumnos de primaria. En este ejemplo el cero es real pues representa que el alumno no ha sido expuesto a los programas de televisión. Al igual que en el de intervalos podemos realizar diferentes procedimiento aritméticos y estadístico.

Una vez identificada la relación entre el nivel de medición y el reactivo o conjunto de preguntas podemos realizar el análisis estadístico pertinente, por ejemplo

Sexo	1= Mujer 2 = Hombre	Estadística descriptiva Categorías de clasificación
Semestre	1 semestre 2 semestre 3 semestre n semestre	Estadística descriptiva Categorías de clasificación
Promedio de calificaciones del último semestre:	0 – 100	Estadística descriptiva e inferencial.

Un ejemplo de una escala para registrar la actitud de un individuo puede tomar la forma siguiente:

- Muy de acuerdo
- De acuerdo
- Indiferente
- En desacuerdo
- Muy en desacuerdo

Existen otras alternativas para elegir la respuesta por ejemplo: (Hernández, 2003).

Muy de acuerdo	De acuerdo	Ni de acuerdo, ni en desacuerdo	En desacuerdo	Muy en desacuerdo
Totalmente de acuerdo	De acuerdo	Neutral	En desacuerdo	Totalmente en desacuerdo
Definitivamente sí	Probablemente sí	Indeciso	Probablemente no	Definitivamente no
Completamente verdadero	Verdadero	Ni falso, ni verdadero	Falso	Completamente falso

Una vez que el instrumento tiene el número de reactivos suficiente con la alternativa de respuesta pertinente, el diseño y formato de la cédula es la adecuada, pasamos a aplicar el instrumento a una muestra piloto para analizar el comportamiento del mismo de acuerdo a los criterios psicométricos que a continuación se describen.

# METODOLOGÍA PARA DETERMINAR LA CALIDAD DE LOS REACTIVOS DE UN INSTRUMENTO

## Paso de la metodología

- I. Análisis del examen aplicado
- II. Elaboración de la hoja de cálculo
- III. Definición de la escala de clasificación de la dificultad del ítem
- IV. Análisis de reactivos: índice de dificultad, índice de discriminación.
- V. Análisis de la confiabilidad de las preguntas de un instrumento:  $r$  de Pearson, Alfa de Cronbach.



## **I. Análisis del examen aplicado.**

En este paso de la metodología es necesario tener en cuenta las siguientes características del instrumento, en el ejemplo que aquí se desarrolla de medición de rendimiento académico:

- Tipo de examen.
- Número de preguntas del examen y temáticas que abordan.
- Tipos de preguntas que contiene el instrumento.
- Clave de calificación y criterio de decisión de éxito o fracaso.
- Resultado obtenido en cada ítem en términos del criterio correcto - incorrecto.
- Calificaciones por temáticas y del examen en general.

## **II. Elaboración de la hoja de cálculo**

Una vez que se obtuvo la información del examen y se codificó cada una de las respuestas del alumno, se procede a su captura en una matriz en la que sus columnas harán referencia a la (s) variable (s) como el folio asignado al participante, nombre del estudiante, sexo y el número de la pregunta. Los renglones o filas se asignan a los estudiantes. En la intersección columna – renglón, corresponde a la clave o el valor asignado al dato demográfico (número del participante, edad, sexo, escuela, etc.) y a la respuesta del alumno como se puede apreciar en el figura 1.

Figura 1. Ejemplo de una matriz para concentrar los datos.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Sujeto	Sexo	PREG. 1	PREG. 2	PREG. 3	PREG. 4	PREG. 5	PREG. 6	PREG. 7	PREG. 8	PREG. 9	PREG. 10	PREG. 11	PREG. 12	PREG. 13	PREG. 14
2	1	2	1	0	0	1	0	1	1	1	1	1	1	1	1	0
3	2	2	1	0	0	1	0	0	1	1	1	1	0	0	0	1
4	3	1	1	0	0	0	0	0	1	1	1	1	1	1	0	1
5	4	2	1	1	0	1	1	0	1	1	1	1	1	0	1	1
6	5	1	1	1	0	1	1	0	1	1	1	1	1	0	0	0
7	6	2	0	0	0	1	0	0	1	1	1	1	1	1	0	1
8	7	1	1	0	0	1	0	0	1	1	1	1	1	0	0	1
9	8	2	1	0	0	1	0	0	1	1	1	1	1	1	1	1
10	9	2	1	0	1	1	0	0	1	1	1	1	1	1	1	1
11	10	2	0	0	0	1	0	0	1	1	1	1	1	0	1	1
12	11	2	1	1	0	0	1	0	1	0	1	1	1	1	0	1
13	12	2	1	0	0	1	0	0	1	0	1	0	1	1	0	1
14	13	2	1	1	0	1	0	0	1	1	1	1	1	1	0	1
15	14	1	1	1	0	1	1	0	1	1	1	1	0	1	0	0
16	15	2	1	0	0	0	0	1	1	0	1	1	1	0	0	1
17	16	2	1	1	0	1	0	0	0	1	1	1	1	1	0	1
18	17	1	1	0	0	1	0	0	1	1	1	1	1	1	0	1
19	18	1	0	0	0	1	0	0	1	1	1	1	1	1	0	1
20	19	2	0	1	1	1	0	0	1	1	1	1	1	0	0	0
21	20	2	1	0	1	1	1	0	1	1	1	1	1	1	1	1
22	21	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1
23	22	2	1	0	0	1	0	0	1	1	1	1	1	1	1	0
24	23	1	0	0	1	0	1	1	1	1	1	1	1	1	1	0
25	24	2	1	0	0	1	0	0	1	1	1	1	1	1	1	1
26	25	2	0	1	0	1	0	0	1	0	1	1	1	1	0	1
27	26	2	1	0	0	0	0	0	1	0	1	1	1	0	0	1
28	27	1	1	1	0	1	0	0	1	1	1	1	1	1	0	1
29	28	1	1	1	0	1	1	1	1	1	1	1	1	0	0	1
30	29	2	1	0	0	1	0	0	1	1	1	1	1	1	0	1
31	30	1	1	0	0	0	1	0	1	1	1	1	1	1	1	1
32	31	1	0	1	0	1	0	0	1	0	1	1	1	1	0	1

Las principales recomendaciones para armar una matriz de datos como se muestra en la figura 1 son las siguientes:

1. Es indispensable que los instrumentos o exámenes tengan un número de folio con el que se va a identificar a cada sujeto al vaciar sus respuestas en la matriz.

2. La primera columna y la primera fila son para indicar los casos y los tratamientos respectivamente. En ocasiones es necesario utilizar más de una columna y más de una fila por el tipo de casos que se va a manejar.
3. indicar el rótulo para cada columna. Este rótulo, también conocido como encabezado de columna, deber ser corto y claro pero que informe acerca del contenido, por ejemplo *Preg01, Suma, Prom.* **Microsoft Excel reconoce los encabezados en varias funciones estadísticas y SPSS las utiliza para identificar los resultados del análisis.**
4. Para visualizar con facilidad una columna o grupo de columnas conviene resaltar cada una de ellas usando la paleta de colores. También es útil engrosar las líneas de división.
5. A cada hoja de cálculo en un libro de Excel se le asigna un nombre (etiqueta) que indica la información que contiene, misma que se incluye como título que encabeza la hoja de impresión. Por ejemplo el nombre del libro “Base Biología” y el nombre de la hoja “Sec. B. Juárez” en la figura 1, se interpreta como “*el archivo que contiene los datos del ejercicio de biología que se aplicó al grupo de la Secundaria Benito Juárez*”.
6. Por último, una recomendación muy importante. Al terminar la captura de los datos en la hoja de cálculo se sugiere que este archivo, sea el documento maestro y duplicarlo asignándole un nombre semejante, por ejemplo de “Base Biología” a “Base Biología 01”, “Base Biología 02” ... así sucesivamente conforme sea conveniente generar un nuevo archivo para trabajar algún otro procesamiento de los datos o graficar. Esta precaución permite conservar la base de datos original, evita que toda la información procesada hasta el momento esté contenida en un solo archivo, además de que su manejo y transportación será más fácil. Sin olvidar claro hacer con frecuencia un

respaldo de nuestra información y guardarla en un lugar adecuado. Nunca sabemos cuándo puede fallar el equipo de cómputo.

### III. Definición de la escala de clasificación de la dificultad del ítem

Para poder comparar los resultados, clasificar los ítems y el examen en su totalidad, atendiendo a los índices de dificultad, es necesario establecer una distribución como la que propone Backhoff, Larrazolo y Rosas (2000):

Tabla 1. Índice de dificultad de los reactivos, ítem, pregunta.

5%	Reactivos fáciles
20%	Medianamente fáciles
50%	Con una dificultad media
20%	Medianamente difíciles
5%	Medianamente fáciles

Esta distribución de la dificultad de los reactivos que conforman un examen responde al principio de la distribución normal del área bajo la curva, misma que asegura una distribución similar de los alumnos que saben más, de los que saben menos y de los regulares (Camacho, 2008, Martínez, 2005).

#### IV. Análisis de reactivos:

##### Índice de dificultad

Kerlinger y Lee (2002) afirman que para incrementar la confiabilidad y validez de una prueba es posible ejecutar un conjunto de procedimientos estadísticos llamados *análisis de reactivos*. Consiste en evaluar cada reactivo de forma separada para determinar si el reactivo es bueno o pobre (Escurra, 2011). En pruebas donde las respuestas se evalúan como correctas e incorrectas (como las pruebas cognitivas o de rendimiento escolar), los reactivos se evalúan en términos de su nivel de dificultad (Macia, 2010).

El índice de dificultad se entiende como la proporción de personas que responden correctamente el reactivo y el número total de personas que toman la prueba. El índice de dificultad tiene valores que van de 0 a 1. Cuando el conjunto de estudiantes responde correctamente casi todos los reactivos de una prueba fácil, los valores del índice de dificultad están muy próximos a 1. Por el contrario, ante una prueba que consista de reactivos demasiado difíciles el conjunto de estudiantes responderá de manera incorrectamente a casi todos los reactivos, los valores del índice de dificultad están muy próximos a 0. En ambos casos la prueba no revela las diferencias de las personas en los conocimientos que evalúa. La mayoría de los creadores de pruebas concuerdan en que los mejores reactivos, en términos de dificultad, son aquellos con valores entre .5 y .7 (Kerlinger & Lee, 2002).

Para calcular el índice de dificultad se utiliza la siguiente fórmula:

$$\text{Dificultad del reactivo} = \frac{\text{Número de personas que responden correctamente el reactivo}}{\text{Número total de personas que toman la prueba}}$$

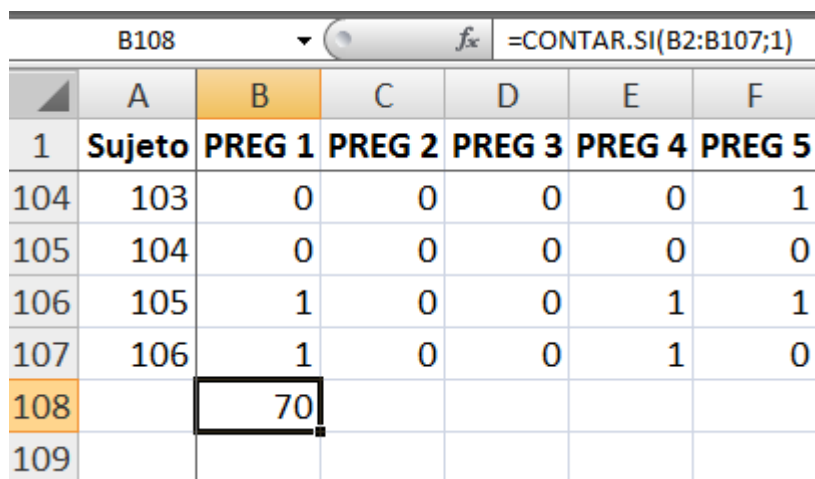
## Desarrollo del ejemplo.

Utilizaremos el archivo *Base Español.xlsx*, que se encuentra en el folder anexo *Bases*.

Paso 1. Por medio de la fórmula =CONTAR.SI(RANGO; CRITERIO) realizaremos el conteo del número de aciertos en la columna PREG 01 (pregunta 01). La sintaxis de la fórmula para esta columna es la siguiente escríbela en la celda B108 =CONTAR.SI(B2:B106;1), (figura 2.)

Paso 2. Cálculo de la Dificultad del Reactivo. En la celda B109 escribe la siguiente fórmula: =B108/106 donde B108 se refiere a la celda que contiene el valor del “Número de personas que responden correctamente el reactivo”, la “/” indica la operación aritmética de dividir y “106” es el “Número total de personas que toman la prueba”.

Figura 2. Conteo del número de aciertos PREG 01.



	A	B	C	D	E	F
1	Sujeto	PREG 1	PREG 2	PREG 3	PREG 4	PREG 5
104	103	0	0	0	0	1
105	104	0	0	0	0	0
106	105	1	0	0	1	1
107	106	1	0	0	1	0
108		70				
109						

Figura 3. Cálculo de la dificultad del reactivo PREG 01.

B109		fx =B108/106			
	A	B	C	D	E
1	<b>Sujeto</b>	<b>PREG 1</b>	<b>PREG 2</b>	<b>PREG 3</b>	<b>PREG 4</b>
104	103	0	0	0	0
105	104	0	0	0	0
106	105	1	0	0	1
107	106	1	0	0	1
108		70			
109		0,66			
110					

Paso 3. Repetir el procedimiento de cálculo para los demás reactivos. Esta acción se realiza de manera muy sencilla con el programa Excel. Primero selecciona las celdas que contienen las fórmulas que deseas repetir en las n celdas contiguas, después coloca el cursor en la esquina inferior derecha de la celda (figura 4). El cursor tomará la forma de una +.

Figura 4. Selección de las celdas para repetir las fórmulas.

B108		fx =CONTAR.SI(B2:B107;1)				
	A	B	C	D	E	F
1	<b>Sujeto</b>	<b>PREG 1</b>	<b>PREG 2</b>	<b>PREG 3</b>	<b>PREG 4</b>	<b>PREG 5</b>
104	103	0	0	0	0	1
105	104	0	0	0	0	0
106	105	1	0	0	1	1
107	106	1	0	0	1	0
108		70				
109		0,66				
110						

Una vez que el cursor toma la forma de + presiona el botón izquierdo del ratón, sin soltar el botón corre el recuadro a la derecha hasta la celda CC109 (PREG 80) observa conforme amplias la selección a la derecha se forma un rectángulo de color gris (figura 5).

Figura 5. Ampliación de las celdas para repetir las fórmulas.

B108		fx =CONTAR.SI(B2:B107;1)					
	A	B	C	D	E	F	G
1	<b>Sujeto</b>	<b>PREG 1</b>	<b>PREG 2</b>	<b>PREG 3</b>	<b>PREG 4</b>	<b>PREG 5</b>	<b>PREG 6</b>
104	103	0	0	0	0	1	0
105	104	0	0	0	0	0	0
106	105	1	0	0	1	1	1
107	106	1	0	0	1	0	0
108		70					
109		0,66					
110							

Una vez que llegas a la columna CC109 suelta el botón izquierdo, observa que se calculan los valores del número de aciertos y el valor del índice de dificultad en cada uno los 79 reactivos restantes de la base de datos (figura 6).

Figura 6. Resultado de repetir las fórmulas.

CJ121		fx					
	A	BY	BZ	CA	CB	CC	CD
1	<b>Sujeto</b>	<b>PREG 76</b>	<b>PREG 77</b>	<b>PREG 78</b>	<b>PREG 79</b>	<b>PREG 80</b>	<b>PUNTAJE</b>
104	103	0	0	0	0	1	26
105	104	1	0	0	0	1	23
106	105	1	0	1	0	0	45
107	106	0	0	1	0	1	30
108		39	19	40	23	57	
109		0,37	0,18	0,38	0,22	0,54	
110							



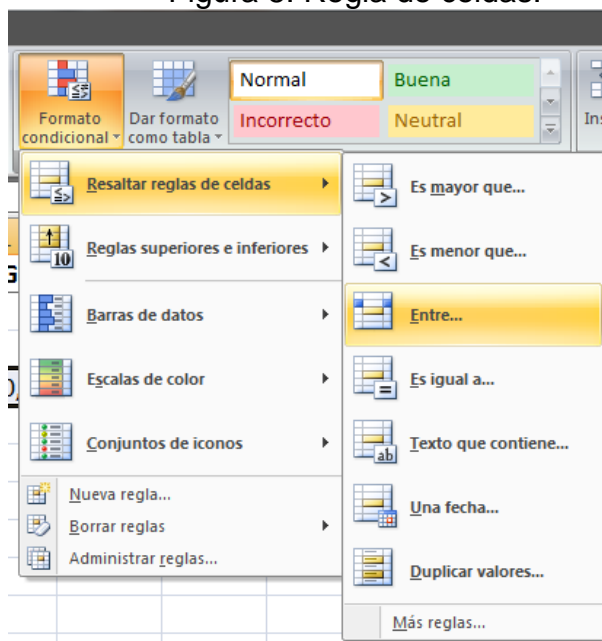
Paso 4. Identificar los *mejores reactivos* con valores entre .5 y .7. Excel ofrece una alternativa para identificar visualmente los mejores reactivos, los reactivos de alta dificultad (por debajo de .49) y los reactivos de baja dificultad (valores superiores a .71). Selecciona el rango B109:CC109 (figura 7).

Figura 7 Selección del rango con la dificultad del reactivo.

	A	BS	BT	BU	BV	BW	BX	BY	BZ	CA	CB	CC
1	Sujeto	PREG 70	PREG 71	PREG 72	PREG 73	PREG 74	PREG 75	PREG 76	PREG 77	PREG 78	PREG 79	PREG 80
104	103	0	0	1	1	0	0	0	0	0	0	1
105	104	0	0	0	1	0	1	1	0	0	0	1
106	105	0	1	0	1	1	1	1	0	1	0	0
107	106	1	1	1	0	0	1	0	0	1	0	1
108		21	30	24	58	44	59	39	19	40	23	57
109		0,20	0,28	0,23	0,55	0,42	0,56	0,37	0,18	0,38	0,22	0,54

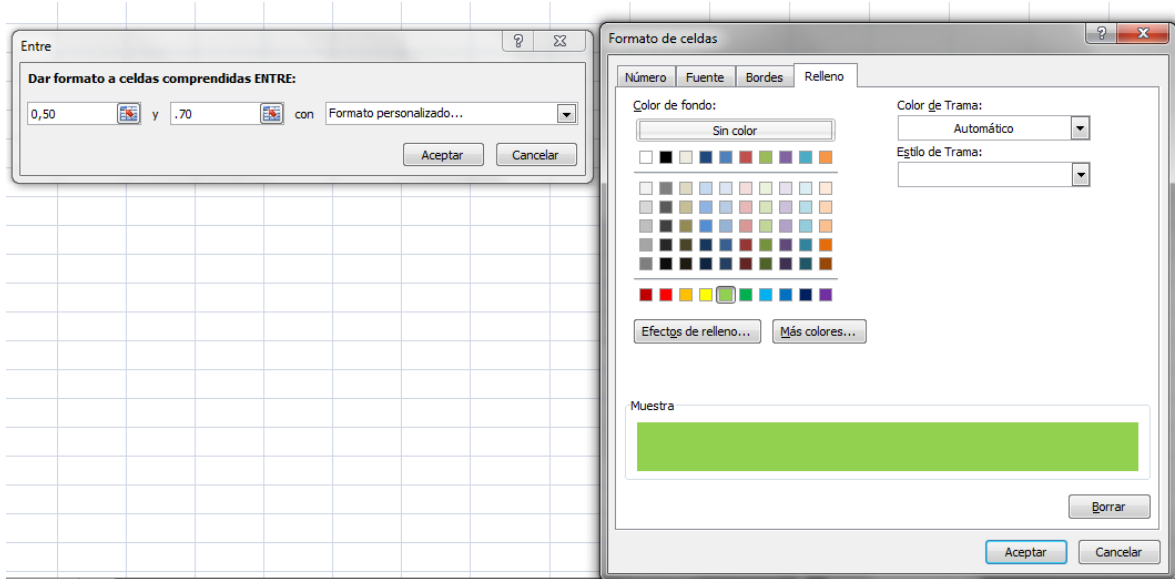
En seguida, en el menú Inicio / Estilo/ Formato condicional/ Resaltar reglas de celdas / Entre, (figura 8).

Figura 8. Regla de celdas.



Selecciona la opción Entre. En la ventana de diálogo Dar formato a celdas compartidas ENTRE, digita los valores .50 y .70 con Formato Personalizado de Celda Relleno color verde. Acepta en ambas ventanas (figura 9).

Figura 9. Formato de celdas.



Las celdas con valor entre 050 a 0.70 tendrán ahora un color verde (figura 10).

Figura 10. Celdas con formato establecido.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	<b>Sujeto</b>	<b>PREG 1</b>	<b>PREG 2</b>	<b>PREG 3</b>	<b>PREG 4</b>	<b>PREG 5</b>	<b>PREG 6</b>	<b>PREG 7</b>	<b>PREG 8</b>	<b>PREG 9</b>	<b>PREG 10</b>	<b>PREG 11</b>	<b>PREG 12</b>
104	103	0	0	0	0	1	0	0	0	1	0	0	0
105	104	0	0	0	0	0	0	0	1	1	0	0	0
106	105	1	0	0	1	1	1	0	1	0	0	1	1
107	106	1	0	0	1	0	0	0	0	1	0	0	1
108		70	24	47	60	53	50	16	76	32	15	38	59
109		0,66	0,23	0,44	0,57	0,50	0,47	0,15	0,72	0,30	0,14	0,36	0,56

## Índice de discriminación

El siguiente índice para el análisis de reactivos es el índice de discriminación de reactivos que indica al investigador (en pruebas cognitivas que tienen respuestas correctas e incorrectas) qué tan efectivamente el reactivo es capaz de discriminar entre puntuaciones altas y puntuaciones bajas (Campo y Oviedo, 2008). Es decir, se considera un buen reactivo aquel que es contestado correctamente por las personas con alta puntuación, y por el contrario, contestado erróneamente por aquellos con baja puntuación (Magnusson, 2009).

Para calcular el índice de discriminación de los reactivos el primer paso es ordenar las puntuaciones totales, de la más alta a la más baja. En seguida, el total de participantes se divide en dos grupos: los de alta puntuación y los de baja puntuación. El criterio de clasificación de ambos grupos varía según el autor; la mayoría coincide en utilizar el percentil 25 ó 27, el percentil 33, inclusive el percentil 50. Es claro que al considerar los extremos (percentil 25), la probabilidad de encontrar mejores índices de discriminación es mayor. También se puede utilizar el criterio del 25 % o la cuarta parte del total de los participantes (Caso, Espinoza, Rodríguez y Urias, 2010).

Valores de 0.0., 1.0 y -1.0 son extraños, en especial el valor negativo que muestra una discriminación invertida (los participantes de menor puntaje –que saben menos- contestan correctamente). Los otros dos valores son interpretados como de una absoluta discriminación (1.0) o bien una nula discriminación (0.0), en ambos casos el reactivo no cumple la función de diferenciar lo que pretende medir el reactivo (Camacho, 2008b). El valor esperado del índice de discriminación tiene que ser positivo y muy próximo a uno; a mayor valor, mayor discriminación (Aiken, 2003).

Para calcular el índice de discriminación se utiliza la siguiente fórmula:

$$\text{Discriminación del reactivo} = \frac{P_A - P_B}{\text{Número de personal en } P_A}$$

Donde:

$P_A$  es el número de personas en el grupo de alta puntuación que respondieron correctamente el reactivo;

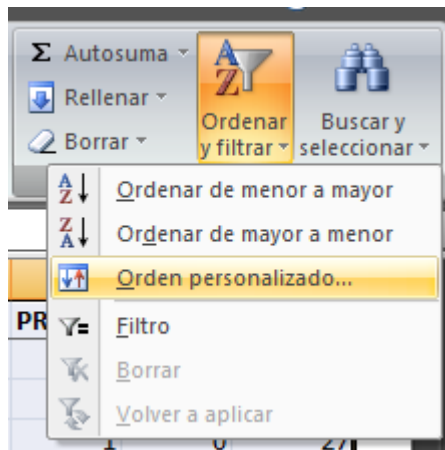
$P_B$  es el número de personas en el grupo de baja puntuación que respondieron correctamente el mismo reactivo;

### **Desarrollo del ejemplo.**

Utilizaremos el archivo *Base Español.xlsx*, que se encuentra en el folder anexo *Bases*.

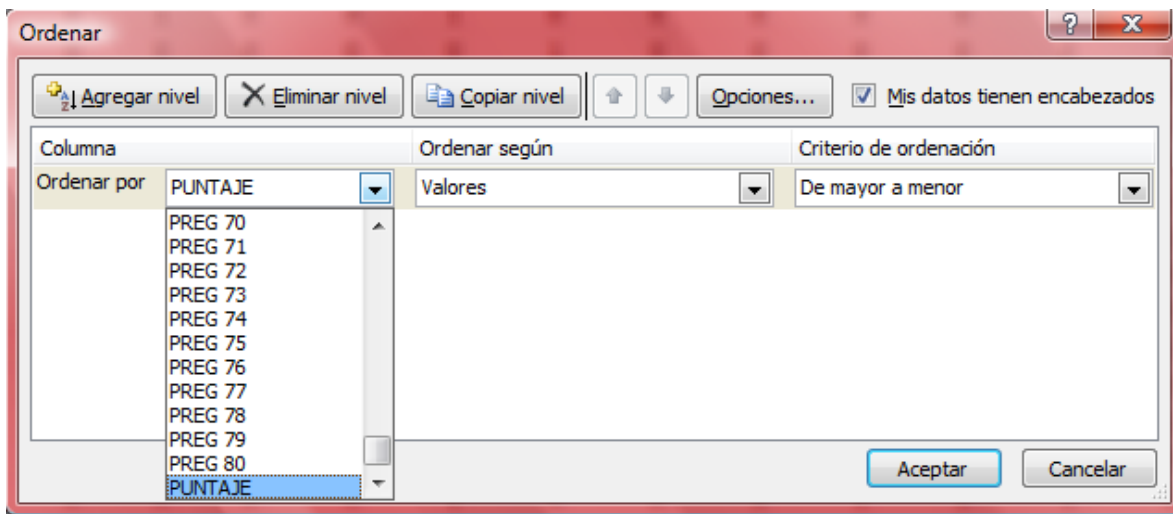
Paso 1. Para mostrar el procedimiento sin modificar el cálculo anterior, realiza una copia de la hoja Base completa y dale el nombre de Discriminación. Recuerda que debes sumar los aciertos de cada alumno como se muestra en la columna Puntaje. Ahora ordena los valores en la columna Puntaje de mayor a menor de la siguiente manera: Selecciona todas las celdas que contienen todos los valores incluyendo los títulos de las columnas; en el menú Modificar, selecciona Ordenar y filtrar / Orden personalizado (figura 11).

Figura 11. Menú Orden Personalizado.



En la ventana que aparece (Ordenar) selecciona en la opción Ordenar por el criterio el valor Puntaje, en la opción Criterio de ordenación selecciona De mayor a menor. Verifica que el cuadro Mis datos tienen encabezados este seleccionado. Da clic en el botón Aceptar (figura 12).

Figura 12. Opciones en la función Ordenar.



Paso 2. Una vez ordenados los valores de la columna Puntaje de mayor a menor, procedemos a identificar la cuarta parte superior e inferior, es decir el 25%. La cuarta parte de 106 es 26.5, por lo que para fines de cálculo el grupo de P<sub>A</sub> (personas de alta puntuación) será de 27 y el grupo de P<sub>B</sub> (personas de baja puntuación) será de 26; esto se debe a que no podemos tomar 0.5 de una persona por lo que ajustamos al entero inmediato superior (26.5 a 27) y al inmediato superior (26.5 a 26).

Contamos 27 personas del puntaje más alto para formar el grupo de P<sub>A</sub> y 27 personas del puntaje más bajo hacia arriba como se muestra en la figura 13.

Figura 13. Ordenar de mayor a menor.

	A	BZ	CA	CB	CC	CD
1	Sujeto	PREG 77	PREG 78	PREG 79	PREG 80	PUNTAJE
8	59	0	0	0	1	43
9	34	0	1	0	1	42
10	62	0	1	0	1	42
11	76	0	1	1	1	42
12	80	0	1	0	1	42
13	6	0	0	0	0	41
14	33	0	0	0	1	41
15	3	1	1	0	1	40
16	53	0	1	0	1	40
17	28	0	1	0	0	39
18	54	0	1	1	1	39
19	73	0	1	0	1	39
20	78	0	0	0	0	39
21	74	0	0	0	1	38
22	99	0	0	1	1	38
23	40	0	1	0	1	37
24	71	0	1	0	1	37
25	93	0	1	0	1	37
26	96	0	0	0	0	37
27	11	0	1	0	0	36
28	77	0	0	0	1	36

Paso 3. Ahora vamos a calcular los valores de  $P_A$  y  $P_B$  para calcular el índice de discriminación de la siguiente manera. En la celda B108 escribe la siguiente fórmula =CONTAR.SI(B2:B28;1) que te permitirá contar el número de aciertos (criterio 1) en el grupo de personas con alta puntuación (rango B2:B28) como se muestra en la figura 14.

Figura 14. Función contar.si PA.

VAR											
=CONTAR.SI(B2:B28,1)											
	A	B	C	D	E	F	G	H	I	J	K
1	Sujeto	PREG 1	PREG 2	PREG 3	PREG 4	PREG 5	PREG 6	PREG 7	PREG 8	PREG 9	PREG 10
100	90	1	0	1	0	0	0	1	0	0	0
101	52	0	0	0	0	1	1	0	0	1	0
102	14	0	0	0	1	0	0	0	1	0	0
103	13	1	0	0	1	1	1	0	0	0	0
104	27	0	1	0	0	1	0	1	1	0	1
105	22	1	0	1	1	0	0	0	0	0	0
106	100	1	0	0	1	0	1	0	1	0	0
107	70	0	0	0	0	0	0	0	0	0	0
108	PA	=CONTAR.SI(B2:B28,1)									
109	PB	CONTAR.SI(rango, criterio)									
110	AltaP										
111	Discrm										

Ahora se calcula el valor de  $P_B$ . En la celda B109 escribe la siguiente fórmula =CONTAR.SI(B82:B107;1) que te permitirá contar el número de aciertos (criterio 1) en el grupo de personas con baja puntuación (rango B82:B107) como se muestra en la figura 15.

Figura 15. Función contar.si PB.

VAR										
	A	B	C	D	E	F	G	H	I	J
1	Sujeto	PREG 1	PREG 2	PREG 3	PREG 4	PREG 5	PREG 6	PREG 7	PREG 8	PREG 9
100	90	1	0	1	0	0	0	1	0	0
101	52	0	0	0	0	1	1	0	0	1
102	14	0	0	0	1	0	0	0	1	0
103	13	1	0	0	1	1	1	0	0	0
104	27	0	1	0	0	1	0	1	1	0
105	22	1	0	1	1	0	0	0	0	0
106	100	1	0	0	1	0	1	0	1	0
107	70	0	0	0	0	0	0	0	0	0
108	PA	20								
109	PB	=CONTAR.SI(B82:B107,1)								
110	AltaP									
111	Discrm									

Paso 4. Ahora se calcula de la siguiente manera el índice de discriminación con los valores que previamente calculamos: En la celda B111 escribe la fórmula  $=(B108-B109)/B110$ . Obtendrás el valor de 0.26. Repite el mismo procedimiento para las demás preguntas como se indicó líneas arriba (figura 16).



Figura 16. Fórmula índice de discriminación.

VAR										
	A	B	C	D	E	F	G	H	I	J
1	Sujeto	PREG 1	PREG 2	PREG 3	PREG 4	PREG 5	PREG 6	PREG 7	PREG 8	PREG 9
100	90	1	0	1	0	0	0	1	0	0
101	52	0	0	0	0	1	1	0	0	1
102	14	0	0	0	1	0	0	0	1	0
103	13	1	0	0	1	1	1	0	0	0
104	27	0	1	0	0	1	0	1	1	0
105	22	1	0	1	1	0	0	0	0	0
106	100	1	0	0	1	0	1	0	1	0
107	70	0	0	0	0	0	0	0	0	0
108	PA	20								
109	PB	13								
110	AltaP	27	27	27	27	27	27	27	27	27
111	Discrm	=(B108-B109)/B110								

Backhoff et al., (2000), retoman la clasificación que Ebel y Frisbie elaboran para determinar la calidad de los reactivos en términos del índice de discriminación. La tabla 2 muestra los valores de discriminación, su correspondiente interpretación y la recomendación para cada uno de estos valores.

Tabla 2. Poder de discriminación de los reactivos según su valor de discriminación (Ebel & Frisbie, 1986, en Backhoff, Larrazolo & Rosas, 2000).

Discriminación	Calidad	Recomendación
> 0.39	Excelente	Conservar
0.39 – 0.39	Buena	Posibilidades de mejorar
0.20 – 0.29	Regular	Necesidad de revisar
0.00 – 0.20	Pobre	Descartar o revisar a profundidad
< -0.01	Pésima	Descartar definitivamente

Entonces como se muestra en la figura 17, el ítem 4 tendría una excelente calidad de discriminación, el ítem 1 se necesita revisar, el ítem 2 se puede descartar o bien revisar y el ítem 7 se descarta definitivamente.

Figura 17. Ejemplo de calidad de discriminación.

A	B	C	D	E	F	G	H	I	J	K
Sujeto	PREG 1	PREG 2	PREG 3	PREG 4	PREG 5	PREG 6	PREG 7	PREG 8	PREG 9	PREG 10
100	1	0	0	1	0	1	0	1	0	0
70	0	0	0	0	0	0	0	0	0	0
PA	20	8	18	19	16	15	4	26	5	2
PB	13	4	8	7	8	8	5	15	9	4
AltaP	27	28	29	30	31	32	33	34	35	36
Discrm	0.26	0.14	0.34	0.40	0.26	0.22	-0.03	0.32	-0.11	-0.06

#### V. Análisis de la confiabilidad de las preguntas de un instrumento: alfa de Cronbach.

Para que los datos obtenidos con un instrumento de medición puedan usarse en situaciones prácticas, estos deben satisfacer ciertas condiciones. Una de esas condiciones es la confiabilidad, término importante en la elaboración y aplicación de un instrumento de medición en psicología. La confiabilidad de un instrumento refiere nociones de estabilidad, fiabilidad y predictibilidad. Un instrumento confiable es aquel cuya medición es consistente, predecible y fiable (Nunnally, 1991).

Magnusson (2009) afirma que los resultados obtenidos con el instrumento en una determinada ocasión, bajo ciertas circunstancias, deben ser los mismos si volvemos a medir el mismo rasgo en condiciones idénticas. Este aspecto en la exactitud de un instrumento de medida es su confiabilidad, es decir se refiere al grado en el que la medición concuerda consigo misma, es la precisión con la que un instrumento de medición mide aquello que

quiere medir e implica la ausencia relativa de errores de medición en el instrumento de medición (Martínez, 2005).

Nunnally (1991) señala que existen varios métodos para calcular la confiabilidad de un instrumento. Los más usados en la construcción de pruebas son la r de Pearson y el alfa de Cronbach, este último es resultado del desarrollo de la pruebas Kuder – Richardson KR20 y KR21. Si bien el alfa de Cronbach fue diseñada para instrumentos con un nivel de medición por intervalos o de razón, varios autores afirman que también puede ser usada en pruebas nominales como las cognitivas donde hay respuestas de acierto – error, con alto grado de eficacia.

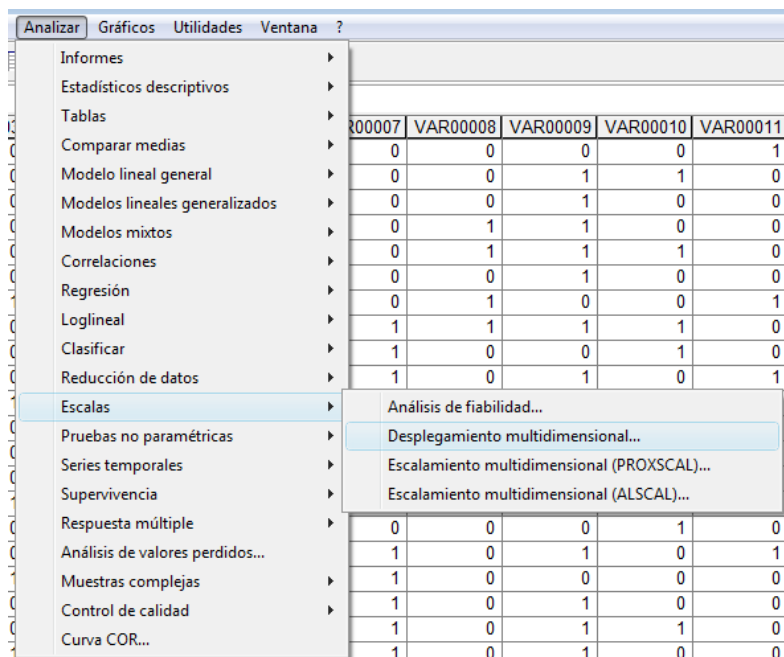
La fórmula para calcular la confiabilidad por medio del alfa de Cronbach es la siguiente:

$$\alpha = \frac{K}{K-1} \left( 1 - \frac{\sum_{i=1}^K S_i^2}{S_T^2} \right)$$

La prueba alfa de Cronbach calcula las varianzas de cada uno de los ítems y la varianza del total de la prueba, lo que la distingue de otros análisis como el de Pearson. Para su cálculo es necesario utilizar el programa SPSS ya que ofrece una mayor agilidad y la posibilidad de obtener más información (Pardo y Ruiz, 2005; Visauta y Martori, 2003).

Paso 1. Abre el archivo Base Español.sav que se encuentra en el folder Español. Del Menú Analizar / Escalas / selecciona Análisis de fiabilidad (figura 18).

Figura 18. Función Análisis de confiabilidad.



En la ventana Análisis de fiabilidad selecciona las 80 preguntas y pásalas a la ventana Elementos como se muestra en la figura 19. Selecciona el Modelo Alfa y en Estadísticos selecciona de la ventana que se muestra en la figura 20 la opción Escala si se elimina elemento.

Figura 19. Ventana de Análisis de confiabilidad, elementos.

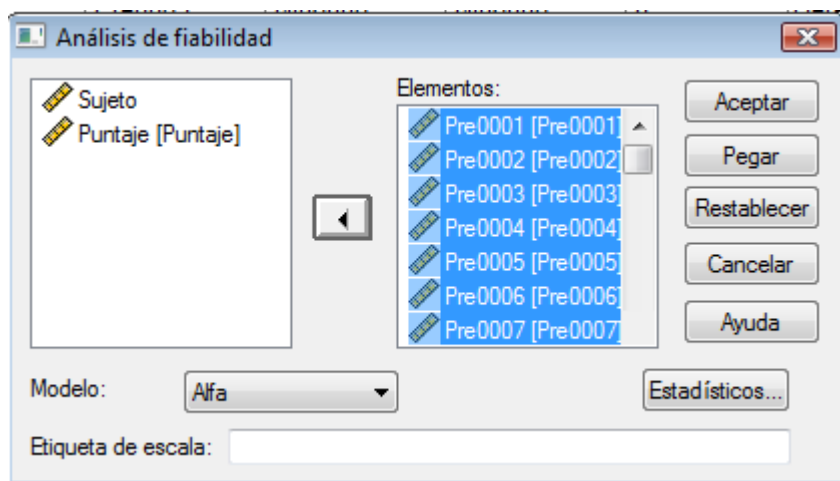
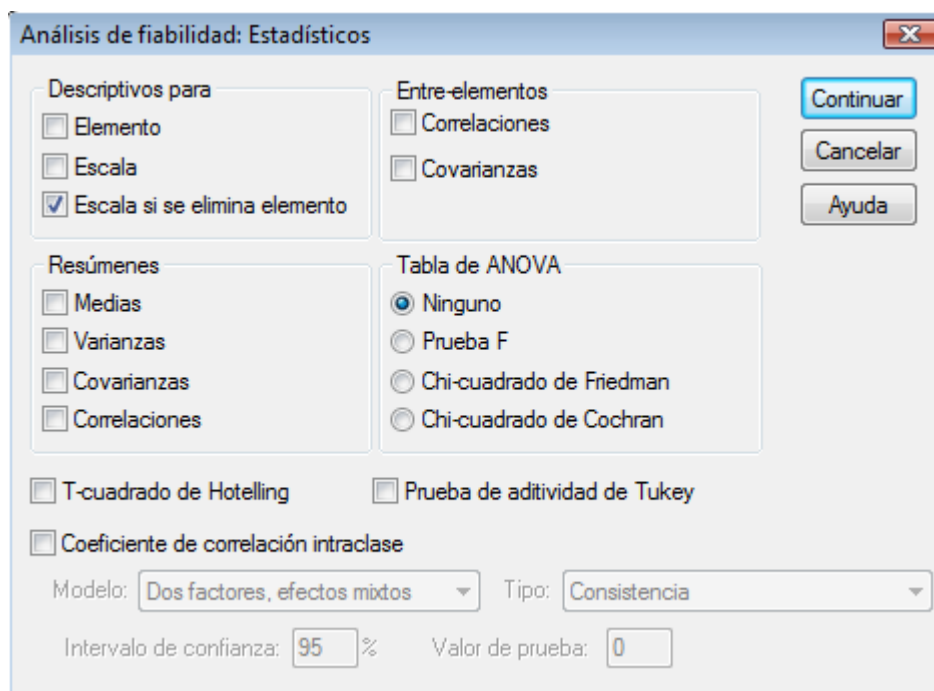


Figura 20. Ventana de Análisis de confiabilidad, estadísticos.



Da clic en el botón aceptar en las ventanas y en la ventana de Resultados podrá ver un conjunto de cuadros. El que corresponde a Estadísticos de fiabilidad (figura 21) contiene el valor para alfa de Cronbach que es 0.707, la podemos considerar aceptable.

Figura 21. Estadísticos de fiabilidad

Alfa de Cronbach	N de elementos
.707	80

El cuadro Estadístico total-elemento, tiene la finalidad de indicar al investigador que sucede con el valor alfa de Cronbach (aumenta o disminuye) sí el ítem se elimina del conjunto (figura 22). Por ejemplo si se retira el ítem 12 el valor de alfa disminuiría a 0.695. Por el contrario sí se retira la pregunta 2 el valor de alfa aumentaría a 0.710. Entre más se

acerque el valor de alfa a 1.0 mejor será el índice de confiabilidad, entre más se acerque el valor a 0.0 la confiabilidad del instrumento será menor.

Figura 22. Estadístico total-elemento.

	Media de la escala si se elimina el elemento	Varianza de la escala si se elimina el elemento	Correlación elemento-total corregida	Alfa de Cronbach si se elimina el elemento
Pre0001	29.95	55.493	.167	.703
Pre0002	30.39	56.868	-.022	.710
Pre0003	30.17	55.171	.200	.702
Pre0004	30.05	55.017	.222	.701
Pre0005	30.11	55.473	.158	.704
Pre0006	30.14	56.027	.084	.707
Pre0007	30.46	57.013	-.043	.710
Pre0008	29.90	54.399	.345	.697
Pre0009	30.31	57.931	-.176	.716
Pre0010	30.47	57.528	-.140	.713
Pre0011	30.25	56.535	.019	.709
Pre0012	30.06	54.016	.360	.695
Pre0013	30.28	54.491	.314	.698
Pre0014	30.25	56.115	.077	.707
Pre0015	30.37	57.149	-.066	.712

Una vez que se realiza el análisis de los reactivos de una prueba de conocimientos, se puede tomar la decisión de cuáles ítems responden a la intención (medición) de nuestro instrumento y cuáles no. En el ejemplo que se viene desarrollando con las preguntas que conforman el módulo de Español, para seleccionar aquellos reactivos que formaran parte de una prueba para el ingreso a la educación media superior (tipo EXANI), se han realizado el análisis del índice de dificultad y de discriminación, así como la confiabilidad global del instrumento. Aquellos reactivos que mejor califiquen a los criterios de estos análisis (como ya se demostró) serán seleccionados para formar parte de la prueba que denominares Ensayo del EXANI- I.

En este momento es oportuno señalar al lector que **realice el ejercicio correspondiente a los 9 módulos restantes** que se encuentran en el paquete de archivos que acompaña este documento. Hay que tener presente que con la ayuda de los programas de computadora, el análisis de las bases de datos se facilita con la automatización o repetición de las funciones.

## **VI. Análisis de los reactivos del EXANI – I**

En el supuesto de que se ha realizado el análisis de los reactivos de cada uno de los diez módulos que forman el EXANI-I, se han elegido las preguntas que mejor califican en el análisis, para ser preciso se han elegido las 12 mejores preguntas de cada módulo y con estas preguntas se elaboró una prueba de Ensayo para el EXANI-I que se ofreció a una escuela secundaria y se aplicó a 48 alumnos.

En este momento el alumno puede realizar el análisis de los reactivos del archivo EXANI-I y poner en práctica los conocimientos adquiridos. El procedimiento es el mismo los criterios de calidad para cada reactivos pueden ser modificados con el fin de observar la manera en que cada ítem puede o no ser considerado de calidad.

Como se puede observar en el archivo de Excel, Ejercicio EXANI-2013, un mayor número de reactivos, en comparación con el análisis original, califican mejor, pero también varias preguntas están fuera de los rangos de consideración. De manera general podemos considerar que el instrumento intitulado en este ejercicio Ensayo para el EXANI-I, es una prueba aceptable en tanto los reactivos que incluye considera preguntas de alto grado de dificultad y de menor grado de dificultad así como preguntas con una dificultad media que la mayoría puede contestar. Además claro de algunas preguntas que muestran valores extraños, por ejemplo el reactivo número 7 con índices de discriminación próximos a cero.

Para concluir, la elaboración de un instrumento que cumpla con los criterios de validez y confiabilidad requiere de un arduo trabajo en su desarrollo y perfeccionamiento, requiere también del conocimiento del conjunto de saberes que se pretende medir, por ello se requiere de la participación de otros especialistas en especial para cumplir con la validez de criterio y de constructo. Además juega un papel muy importante la muestra a la que se le aplica el instrumento pues de acuerdo a sus características (cognitivas, sociales, económicas, etc.) estarán relacionadas sus respuestas a las preguntas. La elaboración de un instrumento válido y confiable para la recolección de datos además de ser una ciencia también es un arte y en este trabajo se mostró una etapa muy importante para su elaboración.



## Bibliografía

1. Adkins, D. (2008). Elaboración de tests. Desarrollo e interpretación de los test de aprovechamiento. México: Trillas.
2. Aiken, L. R. (2003). Test Psicológicos y Evaluación. México: Pearson Educación.
3. Backhoff, E., Larrazolo, N., & y Rosas, M. (2000). Nivel de dificultad y poder de discriminación del Examen de Habilidades y Conocimientos Básicos (EXHCOBA). *Revista Electrónica de Investigación Educativa*, 2 (1). Recuperado de <http://redie.uabc.mx/vol2no1/contenido-backhoff.html>.
4. Baessa, Y. & Arroyave, M. (1996). Confiabilidad y validez del inventario de estrategias de estudio y aprendizaje. *Revista Latinoamericana de Psicología*, 002, 293-303. Recuperado de <http://www.konradlorenz.edu.co/es/publicaciones/revista-latinoamericana-de-psicologia.html>
5. Camacho, J. (a 2008). Asociación entre variables cuantitativas: análisis de correlación. *Acta Médica Costarricense*. 2, 94-96.
6. Camacho, J. (b 2008). Asociación entre variables cuantitativas: correlación no paramétrica. *Acta Médica Costarricense*. 3, 144-146.
7. Campo, A. & Oviedo, H. C. (2008). Propiedades psicométricas de una escala: la consistencia interna. *Revista Salud Pública*. 10, 831-839.
8. Caso, J., Espinoza, L. A., Rodríguez, J. C. & Urias, E. (2010). Propiedades psicométricas de la adaptación del Cuestionario de Estrategias de Aprendizaje para adolescentes. UEE RT 10-001. Ensenada, México. Universidad Autónoma de Baja California, 1-35.
9. Ecurra, L. M. (2011). Análisis psicométrico del cuestionario de Honey y Alonso de Estilos de aprendizaje (CHAEA) con los modelos de la Teoría Clásica de los Test y de Rasch. *Persona*, 14, 71-109.

10. Esteban, M; Ruiz, C., & Cerezo, F. (1996). Validación del cuestionario ILP-R, versión española. *Anales de Psicología*, 12(002), 133-151.
11. Gargallo, B., Suárez, J. & Pérez, C. (2009). El cuestionario CEVEAPEAU. Un instrumento para la evaluación de las estrategias de aprendizaje de los estudiantes universitarios. *Revista electrónica de Investigación y Evaluación Educativa*, 15(2), 1-31. Recuperado de [http://www.uv.es/RELIEVE/v15n2/RELIEVEv15n2\\_5.htm](http://www.uv.es/RELIEVE/v15n2/RELIEVEv15n2_5.htm)
12. Hernández, R., Fernández, C. & Baptista, P. (2003). Metodología de la Investigación. México: McGraw-Hill.
13. Kerlinger, F. N. & Lee, H. B. (2001). *Investigación del Comportamiento. Métodos de investigación en ciencias sociales*. México: Mc GrawHill.
14. Macia, F. (2010). Validez de los test y el análisis factorial: Nociones generales. *Ciencia y Trabajo*. 35, 276-280.
15. Magnusson, D. (2009). *Teoría de los Tests*. México: Trillas.
16. Martínez, R. (2005). *Psicometría: Teoría de los tests psicológicos y educativos*. España: Editorial: Síntesis.
17. Nunnaly, J. L. (1991). *Teoría Psicométrica*. México: Trillas.
18. Pardo, A., y Ruiz, M. A. (2005). Análisis de datos con SPSS 13 Base. España: McGraw Hill.
19. Thorndike, R. L. y Hagen, E. P. (2003). *Medición y evaluación en psicología y educación*. México: Trillas.
20. Tornimbeni, S., Pérez, E., y Olaz, F. (2014). *Introducción a la psicometría*. México: Paidós.
21. Visauta, B. y Martori, J. C. (2003). Análisis estadístico con SPSS para Windows. México: McGraw Hill.
22. Yela, M. (1996). Los test y el análisis factorial. *Psicothema*.8, 73-88.
23. Yela, M. (1996). Los test. *Psicothema*.8, 249-263.

Se encuentran disponibles los archivos en Excel y SPSS para los ejemplos y las prácticas que se indican a lo largo del texto en [juarezlugo@gmail.com](mailto:juarezlugo@gmail.com)