

Construcción y validación de los instrumentos para la medición de la influencia de los campos emocionales en los aprendizajes significativos

Patricio R. Pacheco H., Universidad Tecnológica Metropolitana y U. Tecnológica de Chile (Inacap), Chile

Resumen: Este trabajo presenta un análisis de la confiabilidad y validez de instrumentos de medida desarrollados y empleados para indicar la presencia de conductas específicas, en el tiempo, de alumnos constituidos en equipos que realizaban actividades experimentales en un Laboratorio de Física. Estas conductas están incluidas en las dimensionalidades fundamentales de este estudio $X =$ Indagación / Persuasión asociada a uso de instrumental científico y tecnologías de información y computación, $Y =$ Positividad / Negatividad relacionada a los campos emocionales y $Z =$ Información Interna / Información externa referidas a la clase magistral del docente y al uso de material didáctico. Estas dimensionalidades están interrelacionadas evolucionando según un modelo de dinámica no lineal que da cuenta de la influencia del campo emocional en el proceso. A través del Coeficiente de Correlación de Pearson y el índice de Cronbach se confirma la validez y confiabilidad del instrumental.

Palabras clave: instrumentos de medida, confiabilidad, validez, coeficiente de correlación de Pearson, índice de Cronbach

Abstract: This paper presents an analysis of the reliability and validity of measurement instruments developed and used to indicate the presence of specific behaviors, in time, students in teams that performed experimental activities in a physics lab. These behaviors are included in the fundamental dimensionalities of this study $X =$ inquiry / Persuasion associated with use of scientific instrumentation and information technology and computing, and $Y =$ positivity / negativity related to emotional fields and $Z =$ information internal / external information referred to the master class of the teacher and the use of didactic material. These dimensionalities are inter-related to evolve according to a model of non-linear dynamic that gives account of the influence of the emotional field in the process. Through the Pearson coefficient of correlation and the index Cronbach's alpha is confirmed the validity and reliability of the instruments.

Keywords: Instruments of Measurement, Reliability, Validity, Pearson Correlation Coefficient, Cronbach's Index Alpha

Introducción

Toda medición o instrumento de recolección de los datos debe reunir dos requisitos esenciales: *validez y confiabilidad*.

La validez de un instrumento de medición

La validez

En términos generales, se refiere al grado en que un instrumento realmente mide la variable que pretende medir. Si se desea medir conductas relacionales de un equipo de alumnos debe medir esas conductas.

Aparentemente es sencillo lograr la validez. Parece suficiente, en primera aproximación, razonar la variable de interés y ver cómo construir preguntas sobre esa variable. Esto es factible



en algunos casos (como lo sería el color de piel de una persona). Pero la situación no es tan sencilla cuando se trata de variables como la motivación, la calidad del servicio de un supermercado a los clientes, la actitud hacia un candidato que postula un cargo político y menos aún con sentimientos y emociones, así como diversas variables con las que se trabaja en ciencias sociales. La validez es de complejidad mayor y debe alcanzarse en la aplicación de todo instrumento de medición. Esto se puede plantear a través de la siguiente pregunta: ¿se mide lo que se cree estar midiendo? Si es así, la medida es válida; si no, no lo es (Kerlinger, 1979).

Recomendaciones para aumentar la validez de un diseño de investigación:

Tabla 1. Esquema para la validez de un diseño de investigación

<i>Interna</i>	<i>Externa</i>	<i>De constructo</i>	<i>Estadística</i>
Creación de varios grupos de comparación equivalentes al de observación	Selección de las unidades de la muestra mediante procedimientos aleatorios	Delimitación clara y precisa de los conceptos teóricos	Aumentar el tamaño de la muestra
Efectuar varias mediciones		Operacionalización de los conceptos	Formar grupos internamente poco homogéneos
Controlar todo suceso externo e interno a la investigación que puedan afectar sus resultados	Formar grupos heterogéneos de unidades de observación que incluyen varios contextos temporales y espaciales	Empleo de técnicas de obtención de información	

Fuente: Información adaptada de Kerlinger, 1979, p. 138.

La validez es un concepto del cual pueden tenerse diferentes tipos de evidencia (Wiersma, 1986; Gronlund, 1985) relacionada con el: 1) *el contenido*, 2) *criterio* y 3) *constructo*.

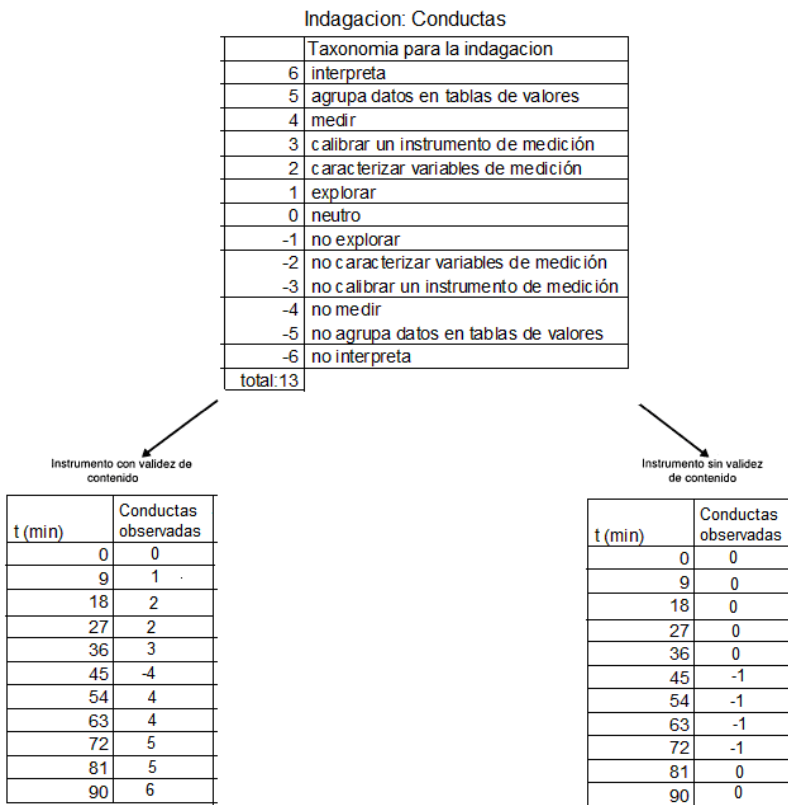
Explayándose respecto de cada una de ellas:

1) Evidencia relacionada con el contenido

La *validez de contenido* se refiere al grado en que un instrumento refleja un dominio específico de contenido de lo que se mide. Es el grado en que la medición representa al concepto medido (Bohrnstedt, 1976). Por ejemplo, una prueba de operaciones aritméticas no tendrá validez de contenido si incluye sólo problemas de resta y excluye problemas de suma, multiplicación o división (Carmines y Zeller, 1979).

Un instrumento de medición debe *contener* a todos los ítems del dominio de contenido de las variables a medir. Este hecho se ilustra en la Tabla N°2.

Tabla 2. El esquema muestra un instrumento con validez de contenido versus uno que carece de este.



Fuente: Información adaptada de Hernández Sampieri, 2006, p. 279.

2) Evidencia relacionada con el criterio

La *validez de criterio* establece la validez de un instrumento de medición comparándola con algún criterio externo. Este criterio es un estándar con el que se juzga la validez del instrumento (Wiersma, 1986). Si los resultados del instrumento de medición se relacionen más al criterio, la validez del criterio será mayor. Por ejemplo, un investigador valida un examen sobre conducción de motos de agua, mostrando la exactitud con que el examen predice lo bien que debe hacerse y, externamente, un grupo de conductores opera una moto de agua siguiendo las rutinas definidas.

3) La *validez de constructo* es probablemente la más importante sobre todo desde una perspectiva científica y se refiere al grado en que una medición se relaciona consistentemente con otras mediciones de acuerdo con hipótesis derivadas teóricamente y que conciernen a los conceptos, variables fundamentales o grados de libertad (o constructos) que están siendo medidos. Un *constructo* es una variable medida y que tiene lugar dentro de una teoría o esquema teórico. A modo de ejemplo, un investigador desea evaluar la *validez de constructo* de una medición particular, tal como una escala de emocionalidad constituida por emociones positivas y negativas de un equipo de alumnos orientados hacia el logro de un objetivo de aprendizaje significativo. Se sostiene que el nivel de emociones positivas hacia una tarea está relacionado positivamente con el grado de persistencia adicional en el desarrollo de la tarea (como suele ocurrir con los empleados con mayor motivación intrínseca y que suelen quedarse más tiempo adicional

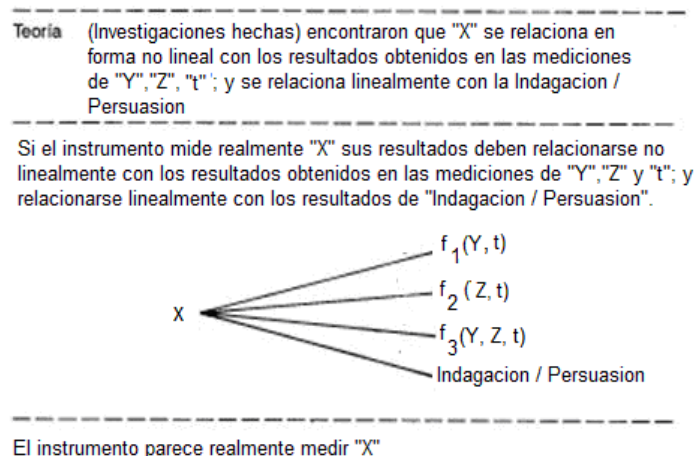
una vez que concluye su jornada). En forma contraria las emociones negativas actúan como un poderoso atractor que atrampa al equipo (Baumeister *et al.*, 2001). Consecuente con esto, la predicción teórica es que a mayor motivación intrínseca, mayor persistencia adicional en la tarea. Si se administra dicho cuestionario de emocionalidad a un grupo de trabajadores también debe determinar su persistencia adicional en el trabajo. Ambas mediciones son correlacionadas. Si la correlación es positiva y sustancial, se aporta evidencia para la validez de constructo.

La validez de constructo incluye tres etapas:

1. Se establece y especifica la relación teórica entre los conceptos (sobre la base del marco teórico).
2. Se correlacionan ambos conceptos y se analiza cuidadosamente la correlación.
3. Se interpreta la evidencia empírica según que tanto clarifica la validez de constructo una medición en particular.

El proceso de *validación de un constructo* está vinculado con la teoría. No es posible llevar a cabo la validación de constructo, a menos que exista un marco teórico que soporte a la variable en relación con otras variables. Desde luego, no es necesaria una teoría sumamente desarrollada, pero sí investigaciones que hayan demostrado que los conceptos están relacionados. Entre más elaborado y comprobado se encuentre el marco teórico que apoya la hipótesis, la validación de constructo puede arrojar mayor luz sobre la validez de un instrumento de medición. Se genera una mayor confianza en la validez de constructo de una medición, si sus resultados se correlacionan significativamente con un mayor número de mediciones de variables que teóricamente y de acuerdo con estudios precedentes están relacionadas. Esto se representa en la Tabla N°3.

Tabla 3. Presentación grafica de un instrumento con validez de constructo
Instrumento mide constructo "X"



Fuente: Información adaptada de Hernández Sampieri, 2006: 283.

También se puede interpretar las evidencias negativas en el marco de la *validez de constructo* (Cronbach y Meehí, 1955; Cronbach, 1984).

La validez total es igual a la validez de contenido más la validez de criterio más la validez de constructo.

Así, la validez de un instrumento de medición se evalúa sobre la base de tres tipos de evidencia. Entre mayor certeza de validez de contenido, validez de criterio y validez de constructo tenga un instrumento de medición, más se acercara a representar la variable o variables que pretende medir.

Un instrumento de medición puede ser confiable pero no necesariamente válido (un aparato —por ejemplo— puede ser consistente en los resultados que produce, pero no medir lo que pretende). Por ello es requisito que el instrumento de medición demuestre ser confiable y válido. De no ser así, los resultados de la investigación no se pueden tomar en serio.

La confiabilidad de un instrumento

La confiabilidad de un instrumento de medición se refiere al grado en que su aplicación repetida al mismo sujeto u objeto, produce iguales resultados. Por ejemplo, al medir en este momento la presión ambiental mediante un manómetro e indicando que hay 1 Atmósfera. Un minuto más tarde se consultara otra vez y el manómetro indicara que hay 6 Atmósferas. Tres minutos después se observara el manómetro y ahora indicara que hay 15 Atmósferas. Este manómetro no sería confiable (su aplicación repetida produce resultados distintos). De la misma manera, si una prueba de habilidades motoras se aplica hoy a un grupo de personas proporcionando ciertos valores y al aplicarse un día después proporcionara valores diferentes, y así en subsecuentes mediciones. Esa prueba no es confiable, suponiendo que el dominio de valores aparece muy extendido. Los resultados no son consistentes; no se puede “confiar” en ellos. La confiabilidad de un instrumento de medición se determina mediante diversas técnicas

Cálculo de la confiabilidad

Existen procedimientos de cálculo formales de la confiabilidad de un instrumento de medición que generan indicadores de confiabilidad. Estos indicadores pueden oscilar entre 0 y 1. Donde un valor 0 significa nula confiabilidad y 1 representa un máximo de confiabilidad (confiabilidad total). Entre más se acerque el valor a cero (0), hay mayor error en la medición. Esto se ilustra en la Tabla N°4.

Tabla 4. Interpretación de un indicador de confiabilidad

	CONFIABILIDAD			
MUY BAJA	BAJA	REGULAR	ACEPTABLE	ELEVADA
0				1
0% de confiabilidad en la medición (la medición esta contaminada de error)				100 % de confiabilidad en la medición (no hay error)

Fuente: Información adaptada de Hernández Sampieri, 2006: 285.

Procedimientos más utilizados para determinar la confiabilidad mediante un coeficiente

1. Medida de estabilidad (confiabilidad por test-retest). En este procedimiento un mismo instrumento de medición (o ítems o indicadores) es aplicado dos o más veces a un mismo grupo de personas, después de un periodo de tiempo. Si la correlación entre los resultados de las diferentes aplicaciones es altamente positiva, el instrumento se considera confiable. Desde luego, el periodo de tiempo entre las mediciones es un factor a considerar. Si el periodo es largo y la variable susceptible de cambios, ello puede confundir la interpretación del indicador o coeficiente de confiabilidad obtenido por este procedimiento. Y si el periodo es corto las personas pueden recordar cómo contestaron en la primera aplicación del instrumento, para aparecer como más consistentes de lo que son en realidad (Bohrnstedt, 1976).

2. Método de formas alternativas o paralelas. En este procedimiento no se administra el mismo instrumento de medición, sino dos o más versiones equivalentes de éste. Las versiones son similares en contenido, instrucciones, duración y otras características. Las versiones —

generalmente dos— son administradas a un mismo grupo de personas dentro de un periodo de tiempo relativamente corto. El instrumento es confiable si la correlación entre los resultados de ambas administraciones es significativamente positiva. Los patrones de respuesta deben variar poco entre las aplicaciones.

3. Método de mitades partidas. Los procedimientos anteriores (medida de estabilidad y método de formas alternas), requieren cuando menos dos administraciones de la medición en el mismo grupo de individuos. En cambio, el método de mitades-partidas requiere sólo una aplicación de la medición. Específicamente, el conjunto total de ítems (o componentes) es dividido en dos mitades y las puntuaciones o resultados de ambas son comparados. Si el instrumento es confiable, las puntuaciones de ambas mitades deben estar fuertemente correlacionadas. Un individuo con baja puntuación en una mitad, tenderá a tener también una baja puntuación en la otra mitad. El procedimiento se presenta en el diagrama de la Tabla N°5.

Tabla 5. Esquema del procedimiento de mitades partidas

El instrumento de medición se aplica a un equipo	Los ítems se dividen en dos mitades (el instrumento se divide en dos)	Cada mitad se califica independientemente	Se correlacionan puntuaciones y se determina la confiabilidad	
<div style="border: 1px solid black; padding: 5px; width: fit-content;"> 1 2 3 4 5 6 7 8 9 10 11 12 </div>	<div style="border: 1px solid black; padding: 5px; width: fit-content; display: inline-block;"> 1 5 8 9 11 12 </div>	resultados (puntuaciones)	P C (0.1) P'	
	<div style="border: 1px solid black; padding: 5px; width: fit-content; display: inline-block;"> 2 3 4 6 7 10 </div>	resultados (puntuaciones)		
	Al dividir los ítems, estos se emparejan en contenido y dificultad.			

Fuente: Información adaptada de Hernández Sampieri, 2006, p. 290.

La confiabilidad varía de acuerdo al número de ítems que incluya el instrumento de medición. Cuantos más ítems la confiabilidad aumenta (desde luego, que se refieran a la misma variable). Esto resulta lógico, al verlo con un ejemplo cotidiano: se desea probar qué tan confiable o consistente es la honestidad de un amigo hacia nuestra persona, cuantas más pruebas le pongamos, su confiabilidad será mayor. Pero demasiados ítems agotaran a la persona que responde.

4. *Coefficiente alfa, α , de Cronbach*. Este coeficiente (Cronbach, 1951) requiere una sola administración del instrumento de medición y produce valores que oscilan entre 0 y 1. Su ventaja reside en que no es necesario dividir en dos mitades a los ítems del instrumento de medición, simplemente se aplica la medición y se calcula el coeficiente. La manera de calcular este coeficiente se mostrara más adelante en el marco de esta investigación.

5. *Coefficiente KR-20* (Kuder y Richardson, 1937). Es un coeficiente para estimar la confiabilidad de una medición, su interpretación es la misma que la del coeficiente alfa.

Mapa para la recolección de datos cuantitativos

El diagrama a continuación es una representación esquemática (Hernández Sampieri *et al.* 2006) de un procedimiento de acopio de datos de mediciones y que ilustra en forma resumida la metodología seguida en este trabajo.

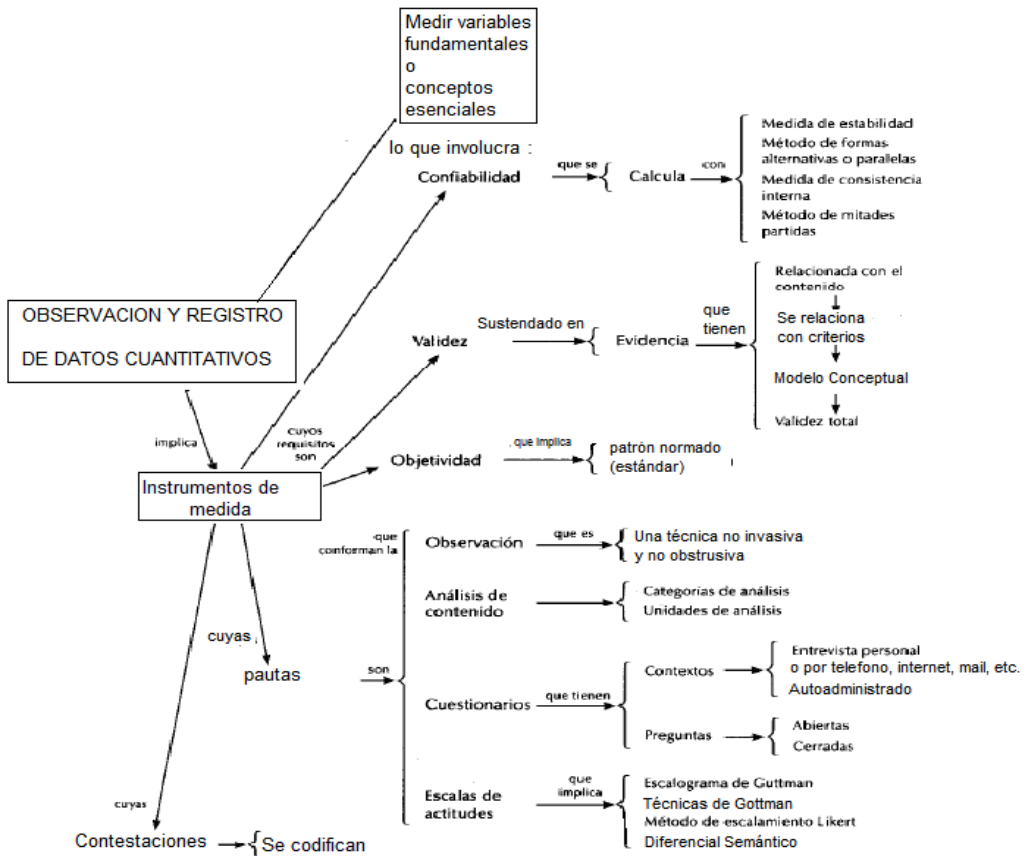


Figura 1: Esquema que muestra un procedimiento de acumulación de datos de mediciones y que resume la metodología seguida en este trabajo
 Fuente(s): Información adaptada de Hernández Sampieri, 2006, p. 308.

Metodología para construir un instrumento de medición

Existen diversos tipos de instrumentos de medición, cada uno con características distintas. Pero, el procedimiento general para construirlos es semejante. Primero es necesario aclarar que en una investigación hay dos opciones respecto al instrumento de medición:

- Elegir un instrumento ya desarrollado y disponible, el cual se adapta a los requerimientos del estudio en particular.
- Construir un nuevo instrumento de medición de acuerdo con la técnica apropiada para ello.

En ambos casos es importante tener evidencia sobre la confiabilidad y validez del instrumento de medición. El procedimiento estándar sugerido para construir un instrumento de medición se presenta a continuación.

Etapas en la construcción de un instrumento de medición

- a) Listar las variables que se pretende medir u observar.

- b) Revisar su definición conceptual y comprender su significado. Por ejemplo, comprender bien qué es la motivación intrínseca y qué dimensiones la integran.
- c) Revisar cómo han sido definidas operacionalmente las variables, esto es, cómo se ha medido cada variable. Ello implica comparar los distintos instrumentos o maneras utilizadas para medir las variables (comparar su confiabilidad, validez, sujetos a los cuales se les aplicó, facilidad de administración, veces que las mediciones han resultado exitosas y posibilidad de uso en el contexto de la investigación).
- d) Elegir el instrumento o los instrumentos (ya desarrollados) que hayan sido favorecidos por la comparación y adaptarlos al contexto de la investigación. En este caso sólo deben seleccionarse instrumentos cuya confiabilidad y validez se reporte. Cualquier investigación sería reportar la confiabilidad y validez de su instrumento de medición. Recuérdese que la primera varía de 0 a 1 y para la segunda se debe mencionar el método utilizado de validación y su interpretación. De no ser así no se puede asegurar que el instrumento sea el adecuado. Si se selecciona un instrumento desarrollado en otro país, deben hacerse pruebas piloto más extensas en la realidad local. También, no debe olvidarse que traducir no es validar un instrumento, por muy buena que sea la traducción. O en caso de que no se elija un instrumento ya desarrollado, sino que se prefiera construir o desarrollar uno propio, debe pensarse en cada variable y sus dimensiones, y en indicadores precisos e ítems para cada dimensión. En este segundo caso, debemos asegurarnos de tener un suficiente número de ítems para medir todas las variables en todas sus dimensiones. Ya sea que se seleccione un instrumento previamente desarrollado y se adapte o bien, se construya uno, éste constituye la versión preliminar de la medición. Versión que debe pulirse y ajustarse en busca del óptimo.
- e) Indicar el nivel de medición de cada ítem y, por ende, el de las variables.

Escalas e índices de confiabilidad

El escalamiento tipo Likert

El instrumento creado para las mediciones de las diferentes dimensionalidades influenciadas por las emociones en el modelo de dinámica no lineal aplica las escalas bipolares.

Ejemplos de escalas Bipolares:

Objeto de estudio: Relaciones al interior de un equipo de alumnos "A"

justas: ____ : ____ : ____ : ____ : ____ : ____ : injustas

Ejemplos de adjetivos bipolares:

	fuerte-débil	poderoso-impotente
grande-pequeño	vivo-muerto	
	bonito-feo	joven-viejo
	alto-bajo	rápido-lento
	claro-oscuro	gigante-enano
	caliente-frío	perfecto-imperfecto
	costoso-barato	agradable-desagradable
	activo-pasivo	bendito-maldito
	seguro-peligroso	arriba-abajo
	bueno-malo	útil-inútil
	dulce-ácido	favorable-desfavorable
	profundo-superficial	agresivo-tímido

Los instrumentos de medición de esta investigación

Variable y = Positividad / Negatividad (Losada y Fredrickson, 2005; Pacheco, Villagrán, Quiroz, 2013).

Es fundamental en este trabajo y corresponde a las emociones (Maturana, 2001; Ibáñez, 2002; Goleman, 2005), que constituyen y dan forma a los campos emocionales, se han dividido en:

Positividad (POS.): se manifiesta en la forma de actos positivos, cuando se muestra apoyo y comprensión dentro de un equipo de trabajo.

Negatividad (NEG.): se manifiesta en la forma de actos negativos, si se demuestra desaprobación, sarcasmo o cinismo

Variable x = Indagación / Persuasión (Losada y Fredrickson, 2005; Pacheco *et al.*, 2013)

Se ha dividido en:

Indagación (IND): se relaciona con utilizar elementos propios de la actividad experimental de la física con el objetivo de explorar y examinar la validez de una posición: la dualidad orientadora es la de “pregunta – medición”, lo que genera descubrimiento.

Persuasión (PER): emplear componentes de discusión a favor del punto de vista de alguien del equipo consiguiendo que los otros comprendan y asuman mi punto de vista.

Variable z = Información interna / Información externa (Losada y Fredrickson, 2005; Pacheco *et al.*, 2013)

Es constituida por:

Información Externa (I.E): este concepto está asociado a si la información recibida orientada a las actividades, persona o grupos, es hacia fuera o desde fuera del equipo y del aula.

Información Interna (I.I): se refiere a la(s) persona(s) que abordan los temas apelando a recursos al interior de un equipo: alumno con habilidades matemáticas, otro de buen manejo de conceptos de física u otro de claras competencias numéricas, etc. Estas actividades fortalecen el análisis de objetivos al interior de un equipo.

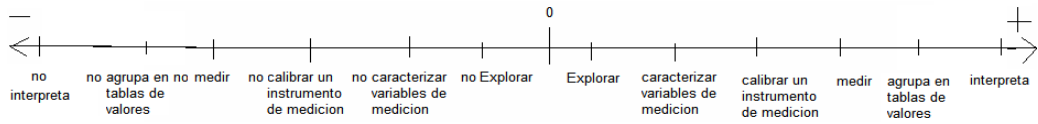
Se realizó seguimiento observacional (Bareman y Gottman, 1989) a un conjunto de 160 alumnos de primer año, ya en su segundo semestre, y que estudian una carrera de ingeniería. Estos alumnos conforman el segmento de alumnado que ingresa solo con sus estudios secundarios terminados sin prerrequisito de pruebas de selección y una proporción importante de ellos, conforman un grupo humano de primera generación en acceder a educación superior (Pacheco *et al.*, 2013).

En total se realizó seguimiento a cuatro cursos de Física Mecánica (nivel introductorio) y a un Laboratorio de Higiene (asignatura de especialidad pero basada en principios de la Física) más un curso de control. Cada curso con un promedio de 32 alumnos que se dividen, para las sesiones específicas de Laboratorio, en dos de 16 alumnos conformándose 4 equipos de trabajo con 4 alumnos cada uno. La matriz de desempeño “típica” por equipo de alumnos está dividida en intervalos de 9 minutos (tiempo basado en criterios de optimización atencional) hasta completar la clase lectiva de 90 minutos.

Cada equipo de alumnos fue observado por tres personas, psicólogos o profesionales capacitados, durante el proceso de seguimiento. Cada una de ellas mide solo una de las dimensionalidades involucradas en este estudio (x, y, z). Es decir se requirieron 12 personas para un total de cuatro equipos cada vez que se realizó la medición. Resultó efectivo, con el propósito de validar una técnica no invasiva de medición, conformar los equipos de seguimiento observacional con alumnos de cursos superiores, previamente seleccionados y capacitados durante un mes por un psicólogo, pues no generan reacciones adversas en los estudiantes bajo observación.

La escala de doble polaridad para la Indagación y la Persuasión y las dos cartillas de seguimiento que controla un observador, especializado en esa variable y que permitía construir la Tabla de Valores que las relaciona, generándose la variable $x(t)$ ($= IND. / PERS. (t)$), y que posteriormente se gráfica:

Indagación:



Esquema N° 1. Señala el dominio de doble polaridad y la escala de codificación, en su doble polaridad, de la Indagación

Fuente: Pacheco, 2013.

Tabla 6. Presenta la conducta a seguir en el tiempo y su taxonomía de codificación.

Codificación	Indagacion Conducta
6 interpreta	0
5 agrupa datos en tablas de valores	9
4 medir	18
3 calibrar un instrumento de medición	27
2 caracterizar variables de medición	36
1 explorar	45
0 neutro	54
-1 no explorar	63
-2 no caracterizar variables de medición	72
-3 no calibrar un instrumento de medición	81
-4 no medir	90
-5 no agrupa datos en tablas de valores	
-6 no interpreta	
total: 13	

Fuente: Pacheco, 2013.

Persuasión:



Esquema 2. Señala el dominio de doble polaridad y la escala de codificación de la persuasión.

Fuente: Pacheco, 2013.

Tabla 7. Presenta la conducta a seguir en el tiempo y su taxonomía de codificación

Codificación		t (min)	Persuasion Conducta
6	Aportar experiencias personales o ajenas que casen bien con nuestra línea argumentativa.	0	
		9	
5	Conectarse directamente con el sistema de creencias del interlocutor.	18	
		27	
		36	
4	Crear un clima de confianza para que la otra persona manifieste sus dudas y dificultades.	45	
		54	
		63	
		72	
		81	
		90	
3	Utilizar la demostración (Video permite ver el desenlace de la historia, sistema multimedia, laboratorios de simulación computacional interactivos, TIC, etc.) ayuda mucho a que se produzca el efecto de imitación		
2	Interlocutor calificado y racional es más argumentativo y eficaz en su conducta persuasiva		
1	Usar estrategias identificativas		
0	neutro		
-1	No 1		
-2	No 2		
-3	No 3		
-4	No 4		
-5	No 5		
-6	No 6		
total:13			

Fuente: Pacheco, 2013.

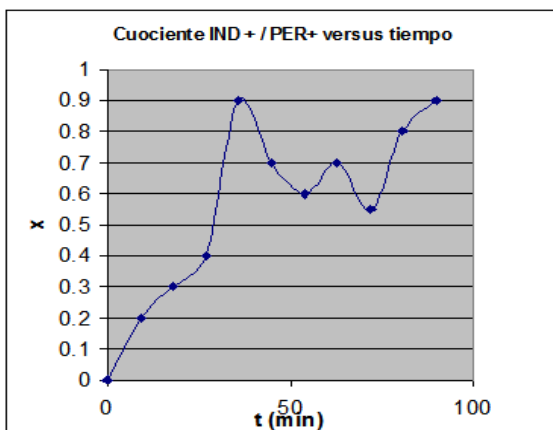


Gráfico 1. La curva es la representación gráfica de la Tabla de Valores del comportamiento en el tiempo, t, de la variable x (definida como cociente Indagación / persuasión).

Fuente: Pacheco, 2013.

Igual tratamiento se realiza para las variables y y z lo que permite, en definitiva, construir las otras Tablas de Valores y sus respectivas representaciones graficas.

Índice alfa, α , de Cronbach

El alfa de Cronbach permite cuantificar el nivel de fiabilidad de una escala de medida para la magnitud inobservable construida a partir de las variables observadas. En psicometría el Alfa de

Cronbach es un coeficiente que sirve para medir la fiabilidad de una escala de medida, y cuya denominación Alfa fue realizada por Cronbach (Cronbach, 1951), aunque sus orígenes se encuentran en los trabajos de Hoyt (Hoyt, 1941) y de Guttman (Guttman, 1945).

Formulación del alfa de Cronbach (Cronbach, 1951; Cronbach y Meehl, 1955)

El alfa de Cronbach, α , es una media ponderada de las correlaciones entre las variables (o ítems) que forman parte de la escala. Puede calcularse de dos formas: a partir de las varianzas (alpha de Cronbach) o de las correlaciones de los ítems (alpha de Cronbach estandarizado). Hay que advertir que ambas fórmulas son dos versiones del mismo concepto y pueden deducirse la una de la otra. El alfa de Cronbach y el alpha de Cronbach estandarizados, coinciden cuando se estandarizan las variables originales (ítems).

A partir de las varianzas

El alfa de Cronbach se calcula de la siguiente forma:

$$\alpha = \left[\frac{k}{k-1} \right] \left[1 - \frac{\sum_{i=1}^k S_i^2}{S_t^2} \right]$$

donde

- S_i^2 es la varianza del ítem i ,
- S_t^2 es la varianza de los valores totales observados los que designamos en forma genérica por y .
- k es el número de preguntas o ítems (cada uno de los apartados que componen un cuestionario o un test).

Varianza caso discreto (Spiegel, 2000)

$\text{Var}(X)$ ($= S^2$) (también representada como $\sigma_{x_0}^2$, simplemente σ^2). Si la variable aleatoria X es discreta con pesos $x_1 \mapsto p_1, \dots, x_n \mapsto p_n$ y n es la cantidad total de datos, entonces tenemos:

$$\text{Var}(X) = \frac{\left(\sum_{i=1}^n p_i (x_i - \mu)^2 \right)}{n}$$

donde:
$$\mu = \frac{\left(\sum_{i=1}^n p_i x_i \right)}{n}$$

A partir de las correlaciones entre los ítems

A partir de las correlaciones entre los ítems, el alfa de Cronbach estandarizado (Cronbach, 1984) se calcula así:

$$\alpha_{est} = \frac{k p}{1 + p(k-1)}$$

donde k es el número de ítems y p es el promedio de las correlaciones lineales entre cada uno de los ítems (se tendrán $[k(k-1)]/2$ pares de correlaciones).

En probabilidad y estadística, la **correlación** indica la fuerza y la dirección de una relación lineal y de proporcionalidad entre dos variables estadísticas. Se considera que dos variables cuantitativas están correlacionadas cuando los valores de una de ellas varían en forma sistemática con respecto a los valores equivalentes de la otra: si se tienen dos variables (A y B) existe correlación si al aumentar los valores de A lo hacen también los de B y viceversa. La correlación entre dos variables no implica, por sí misma, ninguna relación de causalidad

Coefficiente de correlación muestral de Pearson en puntuaciones diferenciales o centradas

Interpretación geométrica

Dados los valores de una muestra de dos variables aleatorias $X(x_1, \dots, x_n)$ e $Y(y_1, \dots, y_n)$, que pueden ser consideradas como vectores en un espacio a n dimensiones, pueden construirse los “vectores centrados” como: $X(x_1 - \bar{x}, \dots, x_n - \bar{x})$ e $Y(y_1 - \bar{y}, \dots, y_n - \bar{y})$. El coseno del ángulo α entre estos vectores es dada por la fórmula para el estimador puntual, r , del coeficiente de correlaciones poblacional, ρ :

$$r = \cos(\alpha) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

A $r = \cos(\alpha)$ se le denomina el coeficiente de correlación muestral de Pearson (Meyer, 1992). El coeficiente de correlación es el coseno entre ambos vectores centrados:

Si $r = 1$, el ángulo $\alpha = 0^\circ$, ambos vectores son colineales (paralelos).

Si $r = 0$, el ángulo $\alpha = 90^\circ$, ambos vectores son ortogonales.

Si $r = -1$, el ángulo $\alpha = 180^\circ$, ambos vectores son colineales de dirección opuesto.

En forma más general: $\alpha = \arccos(r)$. Del punto vista geométrica, no se habla de *correlación lineal*: el coeficiente de correlación tiene siempre un sentido, cualquiera que sea su valor entre -1 y 1. Nos informa de modo preciso, no tanto sobre el grado de dependencia entre las variables, sino más bien de la distancia angular en la hiperesfera n dimensiones. El cuadro de las correlaciones es un método de análisis multidimensional que reposa en esta idea. La correlación lineal se da cuando en una nube de puntos estos se encuentran o se distribuyen alrededor de una recta.

Distribuciones de muestras asociadas al coeficiente de correlación de Pearson.

Para el caso de variables aleatorias con distribución normal bivariada, Fisher (Meyer, 1992; Spiegel, 2000) encontró para una muestra de tamaño n , la distribución muestral de r . Este estimador tiene función de densidad

$$f(r; \rho) = \frac{2^{n-1} (1 - \rho^2)^{\frac{n+1}{2}} (1 - r^2)^{\frac{n-2}{2}}}{(n-1)! \pi} \sum_{i=0}^{\infty} \frac{(2\rho r)^i}{i!} \Gamma^2 \left[\frac{1}{2}(n+i+1) \right]$$

Por medio de transformaciones: $U = \frac{(r - \rho)\sqrt{n}}{1 - \rho^2} \rightarrow Z$ quien se distribuye Normalmente

$N(0,1)$

Es decir, U converge a la variable Z con distribución normal estándar. Cuando $\rho = 0$ la función de densidad del coeficiente de correlación muestral de Pearson es:

$$f(r) = \frac{\left(\frac{n-3}{2}\right)!(1-r^2)^{\frac{n-4}{2}}}{\sqrt{\pi}\left(\frac{n-4}{2}\right)!} \quad -1 \leq r \leq 1$$

Esta función de densidad es similar a la función de densidad de la distribución t de Student. La que se puede obtener por medio de la siguiente transformación:

$$t = r \sqrt{\frac{n-2}{1-r^2}}$$

originando una distribución t de Student con n-2 grados de libertad cuando $\rho = 0$. Si $\rho \neq 0$, la función de densidad de r se estudia con ayuda del estadístico:

$$Z_r = \frac{1}{2} \ln \frac{1+r}{1-r}$$

donde Z_r tiene distribución normal siendo $E(Z_r) = Z_0$ y $V(Z_r) = 1 / (n-3)$. Aquí:

$$Z_\rho = \frac{1}{2} \ln \frac{1+\rho}{1-\rho}$$

Intervalo de confianza

En estadística, se llama **intervalo de confianza** a un par de números entre los cuales se estima cierto valor desconocido con una determinada probabilidad de acierto. Formalmente, estos números determinan un intervalo, que se calcula a partir de datos de una muestra, y el valor desconocido es un parámetro poblacional. La probabilidad de éxito en la estimación se representa con $1 - \alpha$ y se denomina nivel de confianza. En estas circunstancias, α es el llamado error aleatorio o nivel de significación, esto es, una medida de las posibilidades de fallar en la estimación mediante tal intervalo.

En ingeniería y física (Taylor, 1997), el error aleatorio es aquel error inevitable que se produce por eventos únicos imposibles de controlar durante el proceso de medición. Se contraponen al concepto de error sistemático.

En un estudio de investigación, el error aleatorio viene determinado por el hecho de tomar sólo una muestra de una población para realizar inferencias (Meyer, 1992). Puede disminuirse aumentando el tamaño de la muestra. Su procedimiento de cuantificación se sigue de:

1. Test de Hipótesis o Docimasia.
2. Cálculo de intervalo de confianza.

Las fuentes de los errores aleatorios son difíciles de identificar o sus efectos no pueden corregirse del todo. Son numerosos y pequeños pero su acumulación hace que las medidas fluctúen alrededor de una media.

El nivel de confianza y la amplitud del intervalo varían conjuntamente, de forma que un intervalo más amplio tendrá más posibilidades de acierto (mayor nivel de confianza), mientras que para un intervalo más pequeño, que ofrece una estimación más precisa, aumentan sus posibilidades de error. Para la construcción de un determinado intervalo de confianza es necesario conocer la distribución teórica que sigue el parámetro a estimar, que se puede simbolizar por θ . Es habitual que el parámetro presente una distribución normal. También pueden construirse intervalos de confianza con la Desigualdad de Chebyshev.

Intervalos de Confianza derivados de la distribución t de Student

La Distribución t – Student tiene las características generales:

Realiza comparación de medias, es para muestras chicas, es uniparamétrica por cuanto depende solo de los grados de libertad (n) y son más dispersas que la distribución normal.

El procedimiento para el cálculo del intervalo de confianza basado en la t de Student consiste en estimar la desviación típica de los datos S y calcular el error estándar de la media

$$= \frac{S}{\sqrt{n}}, \text{ siendo entonces el intervalo de confianza para la media } = \bar{X} \pm t_{n-1; 1-\frac{\alpha}{2}} \frac{S}{\sqrt{n}} . \text{ Es este}$$

resultado el que se utiliza en el test de Student: puesto que la diferencia de las medias de muestras de dos distribuciones normales se distribuye también normalmente, la distribución t puede usarse para examinar si esa diferencia puede razonablemente suponerse igual a cero. Para efectos prácticos el valor esperado y la varianza son: E(t(n))= 0 y Var (t(n-1)) = n/(n-2) para n > 3

Ilustración del cálculo del coeficiente de correlación muestral de Pearson y del coeficiente Alfa de Cronbach para la variable X = Indagación / Persuasión en el modelo de dinámica no lineal aplicado a las emociones y los aprendizajes

Estudio de fiabilidad de la escala de medición: caso caótico

Dos equipos de diferentes cursos en horarios distintos (sin posibilidad de interferencia) pero la misma asignatura.

Dimensión X = Indagación / Persuasión en el tiempo.

EQUIPO 1 DE 4 ALUMNOS

<i>t (min)</i>	<i>x_i</i>	<i>X_i-0.55</i>	<i>(X_i-0.55)²</i>
0	0	-0.55	0.3025
9	0.2	-0.35	0.1225
18	0.3	-0.25	0.0625
27	0.4	-0.05	0.0025
36	0.9	0.35	0.1225
45	0.7	0.15	0.0225
54	0.6	0.05	0.0025
63	0.7	0.15	0.0225
72	0.55	0	0
81	0.8	0.25	0.0625
90	0.9	0.35	0.1225

N =11	Suma =	6.05	Suma =	0.845
	Suma /11=	6.05 / 11	VAR ₁ =	0.845 / 11
	Prom	= 0.55		= 0.0768

y:

EQUIPO 2 DE 4 ALUMNOS

$t(\text{min})$	x_2	$X_2 - 0.66$	$(X_2 - 0.66)^2$
0	0	-0.66	0.4356
9	0.2	-0.46	0.2116
18	0.5	-0.16	0.0256
27	0.33	-0.33	0.1089
36	1	0.34	0.1156
45	0.75	0.09	0.0081
54	0.66	0	0
63	1	0.34	0.1156
72	0.83	0.17	0.0289
81	1	0.34	0.1156
90	1	0.34	0.1156

N = 11

Suma =	7.27	Suma =	1.2811
Suma / 11 =	7.27 / 11	VAR ₂ =	1.2811 / 11
Prom	= 0.66	=	0.1164

Se desprende:

$(X_1 - 0.55) * (X_2 - 0.66)$
0.1318
0.161
0.04
0.0165
0.119
0.0135
0
0.051
0
0.085
0.119

Suma Total = 0.7368

Coefficiente de correlación muestral de Pearson para la dimensión fundamental, X, referida a 11 ítems de dos cursos diferentes:

$$r = \cos(\alpha) = \frac{\sum_{i=1}^N (x_{1i} - \bar{x}_1)(x_{2i} - \bar{x}_2)}{\sqrt{\sum_{i=1}^N (x_{1i} - \bar{x}_1)^2} \sqrt{\sum_{i=1}^N (x_{2i} - \bar{x}_2)^2}} \approx \frac{0.7368}{\sqrt{0.845} \sqrt{1.2811}} \approx 0.7 \approx p$$

Coefficiente Alfa de Cronbach:

$$\alpha_{est} = \frac{k p}{1 + p(k - 1)} = \frac{11 * 0.7}{1 + 0.7 * (11 - 1)} \approx 0.9625$$

k = número de ítems del instrumento de medición.

Intervalo de confianza al nivel de $\alpha = 0.05$:

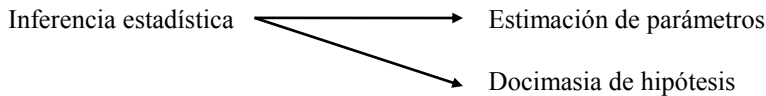
$$\text{Intervalo de confianza} = X \pm t_{n-1; 1-\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$

con $X = 0.55$, $S = 0.2771$, α al nivel 0.05, $r = n - 1 = 11 - 1 = 10 =$ grados de libertad donde n es el numero de mediciones.

$$t_{n-1; 1-\frac{\alpha}{2}} = t_{10-1; 1-\frac{0.05}{2}} = t_{9; 0.975} = 2.228 \text{ (Spiegel, 2000)}$$

$$\text{intervalo de confianza} = 0.55 \pm 2.228 * \frac{0.2771}{\sqrt{11}} = \begin{cases} 0.36 & \text{(mínimo)} \\ 0.74 & \text{(máximo)} \end{cases}$$

Test de Hipótesis o Docimasia (Canavos, 1988; Chao, 1996)



Docimasia: Prueba de Significación Estadística

Probar cierta hipótesis dada para una muestra poblacional.

Etapas de una dócima de hipótesis:

1). Planteamiento de hipótesis

Hipótesis nula	$H_0 :$	=	
Hipótesis alternativa	$H_1 :$	>	Pruebas unilaterales
		<	Pruebas unilaterales
		\neq	Pruebas bilaterales

2) Nivel de significación

Decisión sobre H_0

Realidad de H_0	No rechaza	Rechaza
Verdadero	Correcto	Error I α
Falso	Error II β	Correcto

“ α ” Probabilidad de cometer error tipo I, Rechazar H_0 siendo verdadera

“ β ” Probabilidad de cometer el error tipo II, No rechaza H_0 siendo falsa.

El error tipo I es el nivel de significación de la prueba estadística que siempre ha de indicarse cuando se presentan los resultados.

$\alpha = 0,05$ Rechazar no más de 5 veces en 100 la H_0 siendo V

$\alpha = 0.01$ Rechazar no más de 1 vez en 100 la H_0 siendo V

El error tipo II se denomina potencia o poder estadístico, este depende del tamaño de la muestra ($A >$ tamaño muestral $>$ potencia) y es $= 1 - \beta$.

En los estudios es frecuente establecer como objetivo una potencia de 0.8 con un nivel de significación de 0,05. Ello significa que la probabilidad de cometer un error II (0,2) es cuatro

veces mayor que la probabilidad de cometer un error tipo I (0,05), lo que refleja que en general un error tipo I se considera mucho más grave que un error tipo II.

Aplicación del Test de Hipótesis o Docimasia a la variable de Indagación / Persuasión

Se desea saber si el instrumento de medición esta calibrado desde el punto de vista de la exactitud.

Prueba de Hipótesis:

Se define H_0 = Hipótesis nula y H_1 = Hipótesis alternativa

Procedimiento:

H_0 : $\mu = 0.5$, el instrumento esta calibrado en exactitud.

H_1 : $\mu \neq 0.5$, el instrumento no está calibrado. Hay un error sistemático.

$$t = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} = \frac{0.55 - 0.5}{\frac{0.2771}{\sqrt{11}}} \approx 0.6,$$

donde:

$\bar{X} = 0.55$; $\mu = 0.5$ equipo de control (o validamente calibrado)
 (Para realizar las mediciones y obtener μ se ocuparon equipos de seguimiento observacional diferentes)

$$r = n - 1 = 11 - 1 = 10$$

$$t_{n-1; 1-\frac{\alpha}{2}} = \begin{cases} t_{10; 0.95} = 1.812 \\ t_{10; 0.99} = 2.764 \\ t_{10; 0.999} = 3.169 \end{cases}$$

para los valores numéricos de t (Spiegel,2000)

Representación en el Grafico 2 de la Distribución Normal de los indicadores t:

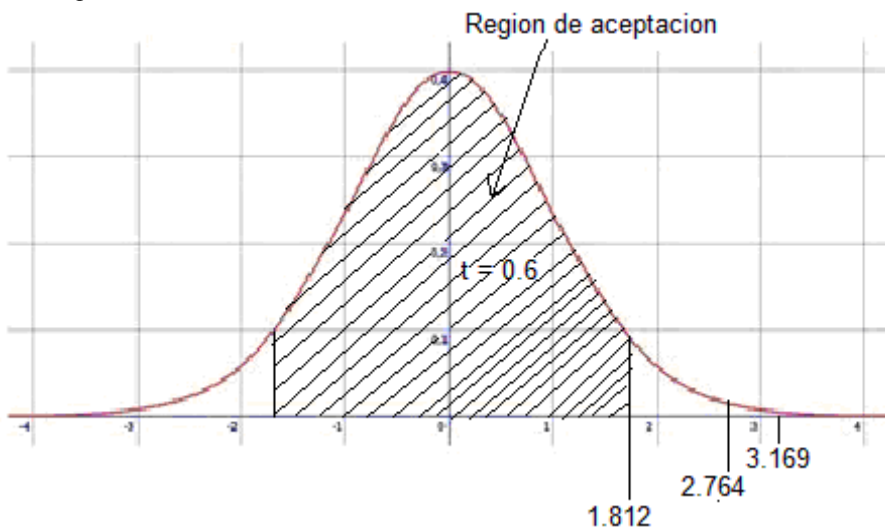


Gráfico 2. Distribución normal de los indicadores t.
 Fuente: Pacheco, 2013.

Se acepta la hipótesis de instrumento calibrado con exactitud.

Realizando un estudio comparativo a partir del coeficiente de correlación muestral de Pearson y del coeficiente alfa de Cronbach de las otras dos dimensiones Y y Z.

Dimensión Y = Positividad / Negatividad en el tiempo

EQUIPO 1 DE 4 ALUMNOS (CURSO 1)

t (min)	Y_1	$Y_1 - 0.63$	$(Y_1 - 0.63)^2$
0	0	-0.63	0.3969
9	0.2	-0.43	0.1849
18	0.3	-0.33	0.1089
27	0.4	-0.23	0.0529
36	0.9	0.27	0.0729
45	0.85	0.22	0.0484
54	0.8	0.17	0.0289
63	0.85	0.22	0.0484
72	0.8	0.17	0.0289
81	0.9	0.27	0.0729
90	0.95	0.32	0.1024

N = 11 Suma = 6.95 Suma = 1.1464
 Suma/11 = 6.95 / 11 VAR1 = 1.1464 / 11
 Prom = 0.63 = 0.104

Se desprende:

EQUIPO 2 DE 4 ALUMNOS (CURSO 2)

t (min)	Y_2	$Y_2 - 0.75$	$(Y_2 - 0.75)^2$
0	-1	-1.75	3.0625
9	-0.5	-1.25	1.5625
18	1.5	0.75	0.5625
27	0.8	0.05	0.0025
36	1	0.25	0.0625
45	1	0.25	0.0625
54	1	0.25	0.0625
63	0.8	0.05	0.0025
72	1.25	0.5	0.25
81	1.2	0.45	0.2025
90	1.2	0.45	0.2025

N=11 Suma = 8.25 Suma = 6.035
 Suma/11= 8.25 / 11 VAR2 = 6.035 / 11
 Prom = 0.75 = 0.5486

donde:

$(Y_1-0.55)*(Y_2-0.66)$
1.1025
0.5375
-0.2475
0.0115
0.0675
0.055
0.0425
0.011
0.085
0.1215
0.144
1.9305

Coefficiente de correlación muestral de Pearson para la dimensión fundamental, Y, referida a dos ítems de dos cursos diferentes:

$$r = \cos(\alpha) = \frac{\sum_{i=1}^N (y_{1i} - \bar{y}_1)(y_{2i} - \bar{y}_2)}{\sqrt{\sum_{i=1}^N (y_{1i} - \bar{y}_1)^2} \sqrt{\sum_{i=1}^N (y_{2i} - \bar{y}_2)^2}} \approx \frac{1.9305}{\sqrt{1.07} \sqrt{2.456}} \approx 0.75 \approx p$$

Coefficiente Alfa de Cronbach:

$$\alpha_{est} = \frac{k p}{1 + p(k - 1)} = \frac{11 * 0.75}{1 + 0.75 * (11 - 1)} \approx 0.97$$

k = número de ítems del instrumento de medición.

Dimensión Z = Información interna / Información externa

EQUIPO 1 DE 4 ALUMNOS (CURSO1)

t (min)	Z ₁	Z ₁ - 0.37	(Z ₁ - 0.37) ²
0	0	-0.37	0
9	0.3	-0.07	0.09
18	0.3	-0.07	0.09
27	0.4	0.03	0.16
36	0.8	0.43	0.64
45	0.3	-0.07	0.09
54	0.3	-0.07	0.09
63	0.7	0.33	0.49
72	0.5	0.13	0.25
81	0.3	-0.07	0.09
90	0.2	-0.17	0.04

N=11

Suma =	4.1	Suma =	2.03
Suma / 11 =	4.1 / 11	VAR1 =	0.5019 / 11
Prom	= 0.37		= 0.0456

Entonces:

EQUIPO 2 DE 4 ALUMNOS (CURSO 2)

t(min)	Z ₂	Z ₂ - 0.565	(Z ₂ - 0.565) ²
0	0	-0.565	0.319225
9	0.33	-0.235	0.055225
18	0.16	-0.405	0.164025
27	0.83	0.265	0.070225
36	0.83	0.265	0.070225
45	0.16	-0.405	0.164025
54	0.83	0.265	0.070225
63	0.75	0.185	0.034225
72	0.83	0.185	0.034225
81	0.83	0.185	0.034225
90	0.66	0.095	0.009025

N = 11 Suma = 6.21 Suma = 1.024875
 prom 6.21 / 11 VAR2 = 1.025 / 11
 = 0.565 = 0.0932

De lo que se sigue:

(Z ₁ - 0.37) * (Z ₂ - 0.112)
0.20905
0.01645
0.02835
0.00795
0.11395
0.02835
-0.01855
0.06105
0.02405
-0.01295
0.01615

0.47385

Coefficiente de correlación muestral de Pearson para la dimensión fundamental, Z, referida a dos ítems de dos cursos diferentes:

$$r = \cos(\alpha) = \frac{\sum_{i=1}^N (z_{1i} - \bar{z}_1)(z_{2i} - \bar{z}_2)}{\sqrt{\sum_{i=1}^N (z_{1i} - \bar{z}_1)^2} \sqrt{\sum_{i=1}^N (z_{2i} - \bar{z}_2)^2}} \approx \frac{0.47385}{\sqrt{0.7084} \sqrt{1.012}} \approx 0.7 \approx p$$

Coefficiente alfa de Cronbach:

$$\alpha_{est} = \frac{k p}{1 + p(k - 1)} = \frac{11 * 0.7}{1 + 0.7 * (11 - 1)} \approx 0.9625$$

k = número de ítems del instrumento de medición.

De los estudios cuantitativos de las mediciones realizadas con los instrumentos implementados se desprende que:

Para la variable X (= Indagación / Persuasión) se acepta la hipótesis de instrumento calibrado con exactitud.

Un tratamiento de iguales características se aplica para las variables Y y Z lo que permite, en definitiva, validar las Tablas de Valores construidas y sus respectivas representaciones gráficas.

Conclusión

Según la base teórica estadística aplicada a los instrumentos de medida y a las técnicas procedimentales, esta nos muestra que se encuentran en rangos de confiabilidad y validación absolutamente aceptables (Taylor, 1997). Si bien el alfa de Cronbach no es un estadístico de uso tradicional, pues no viene acompañado de ningún valor, a priori, que permita rechazar la hipótesis de fiabilidad en la escala. No obstante, cuanto más se aproxime a su valor máximo, 1, mayor es la fiabilidad de la escala. Además, en determinados contextos y por tácito convenio, se considera que valores del alfa superiores a 0,7 o 0,8 (dependiendo de la fuente) son suficientes para garantizar la fiabilidad de la escala.

Los resultados obtenidos son muy prometedores al demostrar que es posible construir, en forma relativamente simple, instrumentos confiables para seguimiento observacional de conductas específicas desplegadas por los diversos equipos de alumnos, formados al interior de una sesión lectiva de 1.5 horas, y que pueden propiciar o no el camino hacia los aprendizajes significativos (Baumeister, Bratslavsky, Finkenauer, 2001). Los equipos han sido motivados (evolución emocional inducida), en la introducción de la clase, a través de condiciones iniciales contextualizadas, para que los procesos de aprendizaje les resulten reveladores, en Ciencias Físicas en particular, lo que en definitiva se confirma en sus dinámicas de trabajo (con atractores de punto fijo, medio o caótico (Pacheco *et al.*, 2013; Pacheco, Villagrán, Guzmán, in press)). Esto abre una puerta, muy promisoriosa, para la puesta a punto, desarrollo y análisis de otras disciplinas de enseñanza por medio de las técnicas estadísticas procedimentales aplicadas, validándose el proceso de medición y registro de datos, lo que en definitiva la da un sustrato riguroso al concepto de aprendizaje significativo.

REFERENCIAS

- Bareman, R., Gottman, J.M. (1989). *Observación de la Interacción: introducción al análisis secuencial*. Madrid: Ediciones Morata S.A.
- Baumeister, R., Bratslavsky, E., Finkenauer, C., and Vohs, K. (2001). Bad is stronger than good. *Review of General Psychology*, 5(4), 323-370.
- Bohrstedt, G.W. (1976). *Evaluación de la confiabilidad y validez en la medición de actitudes*. México D.F.: Editorial Trillas.
- Canavos, G.C. (1988). *Probabilidad y estadística*. Buenos Aires: McGraw-Hill.
- Carmines, E.G. and Zeller, R.A. (1979). Reliability and Validity Assessment. Nº17. Berverly Hills, CA: Sage University Paper.
- Chao, L.L. (1996). *Estadística para las Ciencias Administrativas*. Mexico D.F.: McGraw - Hill.
- Cronbach, Lee J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.
- (1984). *Essentials of psychological testing*. New York: Harper & Row.
- Cronbach, Lee J., and Meehl, Paul E. (1955). Construct Validity in Psychological Tests. *Psychological Bulletin*, 52, 281 – 302.
- Goleman, D. (2005). *La inteligencia emocional*. Barcelona: Kairós.
- Guttman, Louis. (1945). A basis for analyzing test-retest reliability. *Psychometrika*, 10(4), 255-282.
- Gronlund, N. (1985). *Medición y evaluación de la enseñanza*. México, D.F.: Centro Regional de Ayuda Técnica. Agencia para el Desarrollo Internacional.
- Hernández Sampieri, E., Fernández Collado, C., Baptista Lucio, P. (2006). *Metodología de la investigación*. México, D.F.: McGraw-Hill
- Hoyt, C. (1941). Test reliability estimated by analysis of variante. *Psychometrika*, 6(3), p. 153-160.
- Ibáñez, Nolf. (2002). Las emociones en el aula. *Estudios Pedagógicos*, 28, 31- 45.
- Kerlinger, Fred N. (1979). *Enfoque conceptual de la investigación*. Mexico, D.F.: Interamericana.
- Kuder, G.F., Richardson M.W. (1937). The theory of the estimation of test reliability. *Psychometrika*, 2, 151-160.
- Losada, M. y Fredrickson, B.L. (2005). Positive Affect and Complex Dynamics of Human Flourishing. *American Psychologist*, 60(7), 678-86.
- Maturana, H. (2001). *Emociones y lenguaje en educación y política*. Santiago de Chile: Dolmen.
- Meyer, Paul L. (1992). *Probabilidad y aplicaciones estadísticas*. Delaware: Addison-Wesley Iberoamericana.
- Pacheco, P, Villagrán, S, Quiroz, E. (2013). Dinámica no lineal y rendimiento académico: verificación experimental e interpretación. *Revista Internacional de Educación y Aprendizaje*, 1(1), 49 – 73.
- Pacheco, P. Dinámica no lineal y rendimiento académico: verificación experimental e interpretación. Tesis Doctorado en Educación. Universidad Bolivariana de Santiago de Chile, 2013.
- Pacheco H., P., Villagrán R., S., Guzmán A, C. (2015). Estudio del campo emocional en aula y simulación de su evolución durante un proceso de enseñanza-aprendizaje para cursos de Ciencias. *Revista Estudios Pedagógicos* (in press).
- Spiegel, M. (2000). *Estadística*. Mexico, D.F.: McGraw-Hill.
- Taylor, J. (1997). *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*. Sausalito, C.A.: University Science Books.
- Wiersma, W. (1986). *Research methods in education: An introduction*. Boston: Allyn and Bacon.

SOBRE EL AUTOR

Patricio Reinaldo Pacheco Hernández: Licenciado y Profesor de Estado en Física de la Universidad Católica de Valparaíso, Magíster en Física (e) de la Universidad Santiago de Chile, Postítulo en Medio Ambiente de la Universidad Santiago de Chile, Magíster y Doctor en Educación de la Universidad Bolivariana. Académico del Departamento de Física de la Universidad Tecnológica Metropolitana (2014 a la fecha), Docente del Área de Ciencias Básicas de la Universidad Tecnológica de Chile (1993 a la fecha), Docente del Instituto de Matemáticas, Física y Estadística de la Universidad de las Américas (2004 a la fecha). Investigador activo con asistencia a congresos y con publicaciones en el área de Ciencias de la Educación; en particular, en sistemas de alta complejidad, campos emocionales, materiales didácticos y aprendizajes significativos.