

Análisis de un instrumento para monitorear acciones de un programa educativo: Confiabilidad y validez con R.

Recchioni, L. (1)

Analysis of an instrument to monitor the actions of an educational program: reliability and validity with R.

Abstract

The Inclusion and Educational Equality Program (PIIE, in Spanish) has been performed in La Rioja Province from year 2004. The objective of the program is to strengthen the instruction in general basic education schools located in vulnerable areas of the cities by means of five intervention lines. During year 2005, information from a sample of schools was collected and all teachers were interviewed to measure their perception of themselves regarding achievements and difficulties they encountered to perform each line of the program. This research work analyzes the reliability and validity of the instrument construct applied in order to determine if it is qualified enough to be used in the study. The answers of 204 interviewees, which include teachers and managers of educational institutions of the capital city and other cities of the province, were considered. Reliability has been evaluated according to Cronbach coefficient and the validity of the construct through the instrument applied has been evaluated by means of the exploratory factor analysis with varimax rotation. The information has been processed with the reliability analysis available in the Rcmdr package from the R software. The reliability analysis has been performed for each of the five intervention lines and all of the quotients resulted in more than 0.80. Regarding validity, the exploratory factor analysis shows that five extracted factors remove 65% of the items variation or variables included, being the lines considered in the instrument associated to them. The relevance of the items related to the different lines and the need to reformulate one of them has been analysed. It has been concluded that, in general terms, the instrument has the required properties to be used in future applications.

Key words: R, reliability, validity, instrument.

Resumen

En la provincia de La Rioja a partir del año 2004, se ejecutó el Programa de Inclusión e Igualdad Educativa (PIIE) destinado a fortalecer la enseñanza en escuelas de educación general básica en el ámbito urbano ubicadas en zonas vulnerables, mediante cinco líneas de intervención. En el año 2005 se relevó información de una muestra de escuelas y se encuestó a todos los docentes con el fin de medir la percepción de los mismos con respecto a logros y dificultades de la ejecución de cada una de las líneas del programa. En este trabajo se analiza la confiabilidad y validez de constructo del instrumento aplicado, para determinar si el mismo posee las propiedades de calidad que lo habilitan como apropiado para utilizar en el estudio. Se consideraron las respuestas de una muestra de 204 encuestados que incluye docentes y directivos de instituciones educativas de Capital y del interior de la provincia. La confiabilidad se evalúa según el coeficiente de Cronbach y la validez del constructo a través del análisis factorial exploratorio con rotación varimax. El procesamiento de información se realiza con el análisis de fiabilidad que dispone el software R en el paquete Rcmdr. El análisis de confiabilidad se realiza para cada una de las cinco líneas de intervención, lográndose para todas coeficientes mayores de 0,80. Con respecto a la validez, el análisis factorial exploratorio indica que cinco factores extraídos remueven el 65% de la variación de los ítems o variables incluidas, asociándose estos a las líneas consideradas en el instrumento. Se analiza la pertinencia de los ítems relacionados a las diferentes líneas y la necesidad de reformulación de uno de ellos. Se concluye que, en términos generales el instrumento posee las propiedades para su utilización en futuras aplicaciones.

Palabras clave: R, confiabilidad, validez, instrumento.

(1) Universidad Nacional de La Rioja, Av. Carlos Menem y Av. René Favaloro, e-mail: ide.unlar@gmail.com

El Programa Integral para la Igualdad Educativa (PIIE), es una intervención en el sistema educativo que surgió como política educativa nacional en el año 2003 -continuando en el año 2008 a través de acciones de fortalecimiento- desde el Ministerio de Educación de la Nación, con el firme propósito de disminuir la desigualdad educativa que afecta fundamentalmente a los alumnos de instituciones educativas de nivel EGB1 y 2 pertenecientes a los sectores urbanos más empobrecidos o marginales.

Para ejecutar el programa, en la provincia de La Rioja, se constituyó un equipo de asistentes técnicos jurisdiccionales con la coordinación de un referente estableciendo una modalidad de trabajo que consistía en el asesoramiento situado y colaborativo a las instituciones educativas que se incluían en el programa.

Las líneas que definen las acciones y estrategias de intervención en las instituciones son:

1. Promover la formulación de iniciativas pedagógicas escolares (IPE) brindando apoyo pedagógico y económico.
2. Capacitación y Formación Docente.
3. Fortalecimiento del Vínculo de la Escuela con la Comunidad.
4. Provisión de recursos materiales.
5. Refuncionalización de la Infraestructura Escolar.

El Equipo Provincial se constituyó en mayo del año 2004 y el trabajo con las escuelas se inició en agosto de ese mismo año. Al inicio del programa, las escuelas beneficiadas fueron 21, luego, en el año 2005 se incluyeron otras 11 instituciones. Este Programa incluye actualmente a 50 escuelas beneficiarias ubicadas en la Capital e interior de la provincia.

A fines del año 2005 el equipo provincial consideró de interés abordar un monitoreo del programa destinado a relevar la opinión y/o percepción de los docentes y directivos de las instituciones con la finalidad de contar con insumos que den cuenta de la marcha del mismo. Para ese fin se construyó un cuestionario con ítems (afirmaciones o preguntas) clasificados u ordenados según las líneas de intervención correspondientes al programa a las cuales se referenciaban, para ser distribuido en una muestra de instituciones en las cuales los docentes y directivos eran los responsables de responderlo.

Antes de realizar el análisis de una encuesta, es conveniente medir la confiabilidad y validez del instrumento utilizado para recolectar la información, con el fin de evaluar las bondades del mismo, realizar ajustes para futuras aplicaciones y tener elementos que podrían ser de utilidad a la hora del análisis final de la encuesta.

En investigaciones de las ciencias de la educación o psicológicas, cuando se utilizan tests, instrumentos o cuestionarios que contienen ítems cuyas respuestas son medidas con una escala Guttman o Likert, al puntaje total obtenido de la suma de respuestas a cada ítem se lo denomina escala, por lo cual se hace referencia al análisis de la confiabilidad y validez de una escala.

Según Cortada de Kohan (1999), “para entender claramente la diferencia entre validez y confiabilidad digamos por último que la validez de una prueba se refiere a su idoneidad, mientras que la confiabilidad se refiere a su precisión”.

El presente trabajo tiene como finalidad validar el instrumento utilizado para monitorear el programa PIIE, con el fin de indagar si el mismo posee las propiedades que aseguran su capacidad para medir de manera consistente lo que se desea medir y si mide lo que se quiere medir.

Según el Diccionario de la Lengua Española (2001) de la Real Academia Española, establece que medir en su primera acepción, es “comparar una cantidad con su respectiva unidad, con el fin de averiguar cuántas veces la segunda está contenida en la primera”. Estas operaciones son relativamente sencillas en mediciones de alturas, pesos, longitudes, ingresos, etc., relacionadas a las ciencias físicas, biológicas o económicas, donde los instrumentos son estándares. En las ciencias de la educación o del comportamiento el fenómeno a medir no es un objeto tangible sino abstracto, para lo cual el instrumento de medición implica una construcción apropiada para la observación de cada fenómeno: aptitud, sentimiento, opinión, percepción, el clima organizacional, etc.

Sampieri (2006) considera que una medición en este caso es “el proceso de vincular un concepto abstracto con indicadores empíricos”. Los indicadores empíricos son frases, afirmaciones, preguntas o situaciones las que son respondidas según una forma de respuesta: abiertas, dicotómicas o categóricas ordinales; los que en general son denominados ítems o reactivos. Estos pueden agruparse en subgrupos o dimensiones de acuerdo a las características que se desean medir en el estudio. Para que este instrumento, denominado en situaciones de evaluación, test o prueba, sea de calidad es pertinente que posea ciertas características o propiedades. Vieytes (2004) sugiere que “Los instrumentos que utilizamos para medir deben ser, a la vez, confiables (es decir que midan siempre del mismo modo) y válidos (que midan lo que se supone que deben medir)”.

Las propiedades necesarias del instrumento de medición, cualquiera sea el área o disciplina de aplicación, según la teoría clásica de las pruebas son: la confiabilidad y validez, por lo cual el mismo debe ser evaluado para observar estas características y determinar si reúne la capacidad para las mediciones requeridas.

.Estas pruebas son ampliamente aplicadas a estudios psicológicos, educativos y de salud. Como ejemplo, entre muchos: Hirsch, A. (2005); Zegers B., Larrain M., Polaino-Lorente, A. et al (2003); Pérez Ibarra M., Pérez Otero N., et al (2007) y Aguilar-Navarro S., Fuentes-Cantú A. et al (2007).

La teoría clásica de puntuaciones de los test considera que la observación empírica es igual a la observación teórica no observable más un error, que se supone aleatorio. Los errores de medición se dan en todos los campos de la ciencia y cuando las mediciones están libres de error se consideran confiables.

Se entiende por confiabilidad la capacidad del instrumento de medir de forma consistente, precisa y sin error la característica que se desea medir, es decir se espera que el instrumento permita resultados invariantes ante repetidas administraciones al mismo sujeto o bien que posea una consistencia interna en la medición a los sujetos. Abad *et al* (2001) considera que la confiabilidad es el grado de estabilidad, precisión o consistencia que manifiesta el test como instrumento de medición de un rasgo determinado". Es necesario establecer el grado de esa confiabilidad del instrumento y en relación a esto los últimos autores así como Kaplan y Sacuzzo (2006), consideran que la confiabilidad de un test o prueba puede analizarse según:

- La estabilidad temporal,
- Consistencia interna.
- El grado de equivalencia entre dos formas (instrumentos) paralelas.

En el primer criterio debe tenerse en cuenta la prueba Test-Retest que implica dos suministros del mismo instrumento a los mismos sujetos y analizar la correlación entre ambas escalas.

Para evaluar la consistencia interna existen varios métodos: confiabilidad por mitades de los ítems del instrumento, la fórmula Kuder Richarson 20 ó KR_{20} que considera todos los ítems del instrumento aunque con respuestas dicotómicas, el coeficiente de Cronbach que es una adaptación de la anterior para ítems medidos con respuesta ordinal, lo cual implica que es el coeficiente más general para obtener confiabilidad según la consistencia interna. La equivalencia entre formas paralelas implica evaluar a través de una correlación los puntajes o escalas obtenidas con el instrumento a utilizar y otro ya elaborado, que constituya otra versión sobre el fenómeno o característica a medir. Estos últimos autores presentan una explicación muy detallada de los métodos y coeficientes, ya sea orientada al cálculo como a sus alcances.

Cuando se posee un solo instrumento, uno de los coeficientes más utilizados para consistencia interna es el

coeficiente de Cronbach o Alfa de Cronbach que permite evaluar el grado en que diferentes partes del instrumento miden una dimensión de manera consistente, si estuviera diseñado según varias dimensiones, o si la totalidad de ítems lo hacen para una sola dimensión.

El grado de consistencia interna del instrumento o de que diferentes partes (o del mismo miden una dimensión de manera consistente, es posible abordarla a través del coeficiente de Cronbach:

$$\alpha = \frac{k}{k-1} \left(1 - \frac{\sum_{j=1}^k S_j^2}{S_k^2} \right)$$

donde k es la cantidad total de ítems,

$$\sum_{j=1}^k S_j^2$$

es la suma de las variancia de cada ítem y S_k^2 es la variancia de la escala o suma total de los resultados de los k ítems del instrumento, y se lo debe interpretar como la covariación entre los ítems.

Un valor bajo de este coeficiente indicará que los ítems miden constructos diferentes o bien que entre los ítems no existe una consistencia o coherencia. En ítems medidos con escala de Likert cuando los sujetos responden en la opción más alta para algunos lo hacen en la opción más baja para otros. Lo esperable es una consistencia alta, donde los sentidos de las respuestas sean similares, ya que la confiabilidad de la prueba depende del grado, en todos lo ítems miden una dimensión común.

Puede tomar valores negativos e inclusive mayores a 1, teniendo como valor referenciado para la aceptabilidad del instrumento un coeficiente mayor a 0,70, no obstante Kaplan y Saccuzzo sugieren que "para tomar una decisión que afecta el futuro de una persona, los evaluadores deben tratar de encontrar una prueba con una confiabilidad mayor a 0,95".

Si bien la confiabilidad es una propiedad importante en el análisis de los instrumentos de medición, mucho más lo es la validez. Cea D'Ancona (1996) enfatiza que "antes que fiables, los indicadores han de ser válidos. Es decir han de proporcionar una representación adecuada del concepto teórico que miden. Ello es independiente de si reúne o no condiciones de fiabilidad", en ese caso la autora denomina la confiabilidad como fiabilidad, términos que se usan indistintamente en la mayoría de la bibliografía consultada.

Kaplan y Saccuzzo (2006) establecen que 'la validez puede definirse como la concordancia entre la puntuación o medida de la prueba y la cualidad que se cree que está midiendo.

Líneas	Ítem	Variable
Iniciativas Pedagógicas Escolares	La IPE promovió el trabajo en equipo, por ciclo/ áreas/ inter-disciplinas en ejecución de la propuesta.	A1
	La IPE promovió la circulación de la información entre los diferentes actores de la Institución	B1
	La IPE promovió la participación y el compromiso de los distintos actores.	C1
	La IPE requirió la creación de tiempos y espacios para el trabajo institucional.	D1
	La IPE produjo algunos conflictos a nivel institucional.	E1
	Cómo valora los logros de los objetivos planteados en la IPE, en su institución.	F1
	Mejoramiento de la Propuesta Pedagógica de la escuela.	G1
	Uso de los recursos por los docentes en el proceso de enseñanza.	H1
Capacitación y formación docente	Uso de los recursos por los alumnos en el proceso de aprendizaje.	I1
	Revisión de las prácticas de enseñanza.	A2
	Nuevas concepciones y abordajes para el desempeño de la función.	B2
	Diversificación de estrategias de enseñanza.	C2
	Diversificación de las actividades para los alumnos.	D2
Vínculo de la escuela con la comunidad	Fortalecimiento del trabajo en equipo.	E2
	Otros	F2
	Participación de la familia de los niños en actividades escolares.	A3
	Participación de referentes de la comunidad (artistas, escritores, animadores, etc.) en actividades escolares.	B3
	Actividades conjuntas con instituciones que rodean a la escuela (club, centros vecinales, salitas de salud, comisaría).	C3
Recursos Materiales	Revisión de la manera de abordar el vínculo con la comunidad.	D3
	Planteo de nuevas estrategias y metodologías para convocar y vincularse con la comunidad.	E3
	Mejora del trabajo áulico.	A4
	Mejora de las producciones de los niños / niñas.	B4
	Aumento de los niveles de retención de los alumnos.	C4
Refuncionalización de la Infraestructura Escolar	Mejora de la motivación de los docentes frente al trabajo cotidiano.	D4
	Planteo de nuevas estrategias y metodologías para convocar y vincularse con la comunidad.	E4
	Mejoramiento de las condiciones de trabajo institucional.	A5
	Utilización de los nuevos espacios con fines curriculares.	B5
	Utilización de los nuevos espacios con fines extra curriculares.	C5
Aprovechamiento por todos los actores de la institución.	D5	
Incorporación de la nueva herramienta didáctica.	E5	
Actitud positiva frente a las nuevas tecnologías.	F5	

Tabla 1: Líneas de ejecución, ítems y variables de análisis.

Las mismas se encuentran distribuidas en el territorio provincial en 4 departamentos. La muestra la conformaron la totalidad de docentes y directivos que respondieron y enviaron el instrumento al equipo provincial del programa. Se incluyeron 204 docentes y directivos de un total de 8 escuelas que presentaron heterogeneidad en los avances. Por esta razón la muestra no probabilística, puede considerarse representativa de las condiciones imperantes en el universo de instituciones.

Análisis de confiabilidad. Se calcula el coeficiente Alfa de Cronbach para evaluar la consistencia interna de los ítems según cada dimensión del programa educativo. Se analiza también la correlación del ítem con la escala total (exceptuando ese ítem). Los análisis se realizan utilizando R, con el paquete Rcmdr. Asimismo se calculan los intervalos de confianza del 95% a través de bootstrap, usando los paquetes Psy y Boot.

Análisis de validez. Se realiza una indagación de la validez de constructo, mediante un análisis factorial para probar que los ítems corresponden a las cinco líneas operacionalizadas. Este análisis se realiza también con R utilizando una rotación varimax a través del paquete Rcmdr.

Resultados

El análisis de confiabilidad de la primera dimensión, “Promover la formulación de iniciativas pedagógicas escolares (IPE) brindando apoyo pedagógico y económico”, indica que el alfa es adecuado ya que alcanza un valor de 0,81 (Ver ANEXO II). No obstante el ítem E1 que se vincula con la indagación hacia la aparición de “algunos” conflictos institucionales presenta una dificultad ya que eliminándolo, se obtiene un alfa cercano a 0,88.

Por otro lado la correlación entre los valores del ítem con la escala total es aproximadamente igual a 0, lo que implica que el mismo no puntúa en el mismo orden que la escala total. Una de las razones por las cuales este resultado se manifiesta, se relaciona con la construcción del ítem ya que la orientación o dirección de las respuestas es diferentes para puntajes altos. Una recodificación de las respuestas, invirtiéndolas no arrojó un resultado mejor, 0,78.

En relación a la segunda línea programática “Capacitación y formación docente”, la consistencia interna es buena ya que el alfa es cercano al 0,81. La única eliminación posible sería la correspondiente al ítem “Otros” el que no merece una eliminación o reformulación en función de la necesidad de mantenerlo como opción de respuesta.

Los cinco ítems incluidos en la línea “Fortalecimiento del Vínculo de la escuela con la comunidad” le dan a la misma consistencia adecuada con un alfa de 0,87 y un análisis de los coeficientes eliminando alguno de ellos, no mejora la consistencia. Paralelamente las correlaciones entre las puntuaciones individuales y la escala son apropiadas.

Una situación similar se presenta con la línea vinculada con la “Provisión de recursos materiales” aunque el coeficiente alfa para la dimensión es de mayor magnitud, el total alcanza el 0,95. No es apropiado en este caso, eliminar algún ítem ya que son pocos y ninguna eliminación mejora la consistencia.

Una alta consistencia interna se presenta en la última línea “Refuncionalización de la infraestructura escolar” cuyo coeficiente es superior al 0,97, a la vez los ítems no presentan evidencias de la necesidad de eliminar o reformular alguno de ellos.

. Algunas veces la validez se define como la respuesta a la pregunta “¿La prueba mide lo que se supone debe medir?”. Entonces la validez depende fundamentalmente de cómo se han operacionalizado los conceptos a través de los ítems y debe permitir clasificar a los sujetos a los cuales se administra el instrumento.

La validez de un instrumento puede evaluarse a través de una validez de contenido, de constructo y de criterio. La validez de contenido se refiere al grado con el que los ítems cubren los contenidos que querían medirse, es decir, si los ítems cubren todos los conceptos que se querían medir. En la mayoría de los casos solo se extraen algunos ítems de un conjunto mayor, lo cual implicaría saber si los incluidos en el instrumento, constituyen una buena muestra del conjunto mayor. El proceso de validez de contenido es básicamente un análisis cualitativo de contenidos, para el cual pueden utilizarse jueces que valoren la congruencia de los ítems con los conceptos a medir.

La validez de constructo permite analizar si la construcción del instrumento mide las dimensiones propuestas o los conceptos establecidos al inicio del estudio y según los cuales se diseñó el instrumento. Los ítems incluidos en el instrumento deberán hacer referencia a los conceptos establecidos, por lo tanto es necesario probar empíricamente, que ellos se relacionan con las dimensiones establecidas de antemano o bien indagar sobre cuáles son las dimensiones que el instrumento estaría midiendo.

En el primer caso se habla de análisis exploratorio, es decir planteadas al momento del diseño las dimensiones o conceptos. Luego de suministrar el instrumento se realiza un análisis cuantitativo que determine si las dimensiones preestablecidas son las que subyacen en los ítems. Para esto se utiliza una técnica estadística multivariada denominada análisis factorial, con la cual de la matriz de información de los ítems, se extraen factores que representan o engloban subconjuntos de ítems, derivándose luego el análisis, a la observación de que, los factores encontrados son los que se pretendían medir. Habrá ítems que se relacionan más con un factor que con otro, por lo tanto darán información para el factor con el cual tengan una mayor vinculación.

Para detectar qué factores subyacen en la matriz de información, se calculan “cargas factoriales” que son coeficientes de las variables (ítems) originales. A mayores coeficientes mayor aporte posee ese ítem al factor, por este motivo una observación a las cargas factoriales deriva en la denominación o identificación de los factores.

Otra medida de utilidad en este análisis es la unicidad que implica cuánto de la variación del ítem no es explicado por los factores detectados, lo cual indica que ese ítem poco aporta a los factores subyacentes. Por otro lado la validez del constructo exploratorio permite la exploración de las dimensiones o los factores que el instrumento mide cuando no se definieron de antemano.

La validez de criterio implica correlacionar las puntuaciones de la escala, es decir el total de puntuaciones obtenidas a través de todos los ítems con otro criterio o instrumento ya conocido y probado, para determinar si el que se pretende utilizar sigue el mismo patrón de puntajes que el ya conocido.

El objetivo del trabajo es determinar la confiabilidad y validez de un instrumento para la medición de la percepción docente sobre la ejecución del Programa Integral para la Igualdad Educativa.

Materiales y Métodos

Se elaboró un instrumento semiestructurado destinado a directivos y docentes, quienes voluntariamente y de manera anónima respondían al mismo, sobre aspectos relacionados al avance e implementación del programa en su institución. A partir de las cinco líneas de intervención:

1. Promover la formulación de iniciativas pedagógicas escolares (IPE) brindando apoyo pedagógico y económico.
2. Capacitación y Formación Docente.
3. Fortalecimiento del Vínculo de la Escuela con la Comunidad.
4. Provisión de recursos materiales.
5. Refuncionalización de la Infraestructura Escolar,

y agrupadas según estas últimas, se construyeron preguntas abiertas, preguntas medidas con escala de Guttman (Sí o No), y afirmaciones (o preguntas) cuyas valoraciones (o respuestas) se medían con una escala Likert, la cual en todos los casos consideró las opciones: Mucho, moderado, poco o nada y No sabe/No contesta.

Las variables consideradas para el análisis de validación del instrumento fueron los diferentes ítems medidas en escala de Likert y agrupados según las líneas de ejecución, las que se consideran dimensiones del instrumento.

Todos los ítems seleccionados para el análisis de confiabilidad y validez fueron medidos mediante una escala de Likert con las 4 opciones ya presentadas en el apartado anterior, y representan casi la totalidad de ítems considerados en el instrumento, debido a que se descartan sólo dos medidos en una escala de Guttman.

Muestra. La distribución del instrumento se realizó al universo constituido por todas las instituciones del territorio provincial que al 2005 habían implementado el programa PIIE, mediante un envío por correo postal. De la totalidad de instituciones educativas fueron analizadas 8, las que representan a juicio del equipo provincial del programa, las distintas tipologías de escuelas en relación a la apertura, predisposición y logros percibidos por los técnicos.

Posteriormente se calcularon intervalos de 95% de confianza a través de bootstrap y los resultados son los siguientes:

Líneas	Alfa	LI(%2.5)	LS(%97.5)
1	0,8144	0,7628	0,8536
2	0,8074	0,7424	0,8526
3	0,8707	0,8229	0,9027
4	0,9544	0,9318	0,9699
5	0,9750	0,9678	0,9805

Tabla 2: Coeficientes de Cronbach y límites de confianza.

Para analizar la validez del instrumento, además de la validez de contenido realizada a través de la consulta de técnicos del área de educación, se realizó un análisis de validez de constructo mediante un factorial exploratorio, lo que implicó indagar sobre los ítems que se incluyen en los 5 factores asociados a las líneas de indagación del instrumento. El análisis implicó analizar qué ítems se incluyen o resumen las 5 líneas determinadas de antemano, lo cual permite detectar que los propuestos describen o responden a la línea considerada.

Las cargas factoriales (Loadings) presentadas en el ANEXO II para los 5 factores seleccionados con una variancia explicada del 65%, indican que el primer factor corresponde a la línea 5 vinculada con el refuncionamiento de la infraestructura escolar, ya que los coeficientes de los ítems de esa línea superan los valores de 0,87. Sin embargo un ítem relacionado a la línea 1, el H1 también posee un alto coeficiente. Se calculó un coeficiente de Cronbach para los ítems de la línea 5 incluyendo este último ítem y el resultado no fue superador al encontrado con los ítems originales 0,95.

El segundo factor corresponde a la línea 1, aunque en este caso el ítem E1 presenta el menor aporte, a la vez que el H1 presenta la particularidad analizada anteriormente. Por otro lado el ítem E1 no aporta información en ninguno de los factores analizados exceptuando el Factor 5, para el cual su aporte no es muy significativo, 0,382. Por lo tanto ambos ítems debe quedar ubicados en esa línea.

El tercer y cuarto factor están relacionados a la línea 4 y 2 respectivamente, aunque en este último uno de los ítems posee un coeficiente muy bajo, situación que se origina por la naturaleza de la proposición ("Otros"), a la cual sólo se accede en el caso que los anteriores ítems no satisfagan una respuesta. La línea 3 agrupa sus ítems en el quinto factor, aunque aquí también se advierte la presencia de un ítem de esa línea, el E3 con un coeficiente menor en este factor en relación al factor 2. Por lo tanto se calcula el coeficiente de Cronbach incorporando este ítem y el resultado no implica la inclusión del mismo en la línea 2 debido a que el resultado de la consistencia interna no presenta mejoras, 0,82.

Por otro lado, se evaluaron los resultados de unicidad de los ítems considerados que es la variancia residual, es decir la no explicada por los factores.

Los resultados indican que los ítems F2, E1, E2 y D2 son los que menos aportan a la estructura de factores seleccionada, ya que poseen mayores unicidades.

A1	B1	C1	D1	E1	F1	G1	H1	I1	A2	B2	C2	D2	E2	F2	A3
0.439	0.370	0.448	0.538	0.788	0.554	0.484	0.409	0.432	0.268	0.115	0.258	0.627	0.649	0.954	0.368
B3	C3	D3	E3	A4	B4	C4	D4	A5	B5	C5	D5	E5	F5		
0.351	0.316	0.394	0.444	0.141	0.055	0.218	0.173	0.133	0.089	0.145	0.140	0.109	0.093		

El test máximo verosímil de bondad de ajuste para los 5 factores se rechaza, por lo cual se propone analizar las unicidades modelando la información con un ajuste significativo de 14 factores, aunque en la práctica disponer de esa cantidad no aporta a una buena interpretación de la naturaleza de la estructura que subyace en los datos.

Las unicidades siguientes, evidencian que aún considerando 14 factores, continúan siendo importantes la correspondientes a los ítems F2, E1, E2 y D2 lo cual dice que aún considerando un máximo de factores estos ítems no estarían bien redactados ya que es imposible explicar su variación.

A1	B1	C1	D1	E1	F1	G1	H1	I1	A2	B2	C2	D2	E2	F2	A3
0.326	0.221	0.005	0.472	0.630	0.504	0.241	0.097	0.153	0.204	0.035	0.214	0.432	0.519	0.677	0.288
B3	C3	D3	E3	A4	B4	C4	D4	A5	B5	C5	D5	E5	F5		
0.005	0.159	0.264	0.005	0.096	0.056	0.142	0.155	0.085	0.032	0.106	0.005	0.046	0.029		

Conclusiones

Con respecto a la confiabilidad del instrumento se concluye que es confiable ya que tal como se presenta en la Tabla 2, todas las líneas abordadas acusan un coeficiente mayor de 0,70.

Un análisis interno a cada línea indica que en la primera, el ítem E1 no presenta una apropiada formulación y que eliminándolo se mejora la confiabilidad.

En las demás líneas los coeficientes son apropiados y el análisis de eliminación de alguno de ellos no las mejora sustancialmente, con excepción del perteneciente a "Otros" en la línea 2, el que corresponde mantenerlo en el instrumento para permitir la elección de opciones no preestablecidas.

En relación al análisis de validez, se detectan claramente las cinco líneas de abordaje propuestas en el instrumento y aunque algunos ítems presentan una baja carga factorial, del estudio realizado en sus movimientos a otras calculando simultáneamente la confiabilidad, se concluye que los ítems seleccionados al inicio de la investigación no deben trasladarse sino que es necesario mantenerlos tal como están incorporados a cada línea. Por lo tanto, se puede afirmar que los factores que subyacen, en general en los 31 ítems considerados son las dimensiones propuestas y definidas al inicio de la construcción del instrumento, con un 65% de variancia explicada por estos factores.

ANEXO

1. Resultados de los análisis de confiabilidad con Remdr.

Promover iniciativas pedagógicas escolares.

Alpha reliability = 0.8144

Standardized alpha = 0.8379

Reliability deleting each item in turn:

	Alpha	Std.Alpha	r(item, total)
A1	0.7833	0.8119	0.6349
B1	0.7827	0.8111	0.6191
C1	0.7778	0.8064	0.6797
D1	0.7921	0.8200	0.5453
E1	0.8811	0.8840	-0.0936
F1	0.7851	0.8115	0.6293
G1	0.7903	0.8181	0.5623
H1	0.7739	0.8048	0.6848
I1	0.7745	0.8055	0.6786

Capacitación y formación docente.

Alpha reliability = 0.8074

Standardized alpha = 0.8126

Reliability deleting each item in turn:

	Alpha	Std.Alpha	r(item, total)
A2	0.7408	0.7447	0.7309
B2	0.7492	0.7508	0.7058
C2	0.7414	0.7450	0.7343
D2	0.7658	0.7732	0.6169
E2	0.7832	0.7902	0.5434
F2	0.8610	0.8684	0.1557

Vínculo escuela y comunidad.

Alpha reliability = 0.8707

Standardized alpha = 0.8711

Reliability deleting each item in turn:

	Alpha	Std.Alpha	r(item, total)
A3	0.8388	0.8385	0.7185
B3	0.8486	0.8483	0.6746
C3	0.8315	0.8321	0.7428
D3	0.8377	0.8391	0.7186
E3	0.8595	0.8608	0.6299

Recursos Materiales.

Alpha reliability = 0.9544

Standardized alpha = 0.9544

Reliability deleting each item in turn:

	Alpha	Std.Alpha	r(item, total)
A4	0.9375	0.9376	0.8968
B4	0.9246	0.9247	0.9390
C4	0.9502	0.9503	0.8540
D4	0.9472	0.9471	0.8651

Refuncionalización de la infraestructura escolar.

Alpha reliability = 0.975

Standardized alpha = 0.9753

Reliability deleting each item in turn:

	Alpha	Std.Alpha	r(item, total)
A5	0.9706	0.9710	0.9120
B5	0.9678	0.9679	0.9425
C5	0.9724	0.9727	0.8942
D5	0.9705	0.9709	0.9132
E5	0.9701	0.9705	0.9186
F5	0.9693	0.9697	0.9263

Como E1 tiene alto coeficiente en el factor 5 que corresponde al vínculo con la comunidad, calculo el alpha para la línea 3 con el E1:

2. Sentencias en R.

```
> linea1<-read.dbf("linea1")
```

```
> cronbach(linea1)
$sample.size
[1] 204
$number.of.items
[1] 9
$alpha
[1] 0.814438
> cronbach.boot<-function(linea1,x){cronbach(linea1[x,])[[3]]}
> res<-boot(linea1,cronbach.boot,1000)
> quantile(res$t,c(0.025,0.975))
 2.5% 97.5%
0.7627756 0.8536200
```

Loadings:

	Factor1	Factor2	Factor3	Factor4	Factor5
A1		0.702			0.239
B1		0.768		0.117	-0.120
C1		0.714		0.146	0.127
D1		0.651			0.179
E1		-0.124	-0.215		0.382
F1	0.109	0.594	0.128	0.230	0.107
G1	0.138	0.479	0.327	0.389	
H1	0.608	0.364	0.243	0.145	
I1		0.616	0.343	0.215	0.147
A2		0.103		0.815	0.207
B2			0.114	0.926	
C2	0.107	0.317		0.776	0.161
D2		0.203	0.103	0.550	0.137
E2		0.359	0.187	0.387	0.167
F2	-0.116	0.105		0.105	
A3		0.258		0.175	0.723
B3		0.215		0.251	0.729
C3	0.116	0.299		0.149	0.743
D3		0.506	0.109		0.570
E3		0.530	0.223	0.221	0.420
A4	0.217	0.139	0.884	0.107	
B4	0.186	0.173	0.934		
C4	0.130	0.214	0.838	0.110	
D4	0.240		0.862	0.164	
A5	0.908		0.123	0.103	0.111
B5	0.930	0.125	0.143		
C5	0.877	0.240	0.139		
D5	0.908	0.128	0.122		
E5	0.922		0.166		0.109
F5	0.932		0.165		

	Factor1	Factor2	Factor3	Factor4	Factor5
SS loadings	5.302	4.586	3.803	3.193	2.615
Proportion Var	0.177	0.153	0.127	0.106	0.087
Cumulative Var	0.177	0.330	0.456	0.563	0.650

Agradecimientos

Se agradece a los técnicos del Ministerio de Educación de la Provincia de La Rioja, Argentina quienes elaboraron el instrumento y realizaron la recolección de datos.

Referencias

Aguilar Navarro, S.; Fuentes Cantú, A.; Ávila Funes, J. y E. García Mayo (2007). Validez y confiabilidad del cuestionario del ENASEM para la depresión en adultos mayores. *Salud Pública de México / Vol.49, No.4, Julio de 2007.*

Aron A. y E. Aron (2001). *Estadística para psicología.* Pearson Education.

Cea D'Ancona (1996). *Metodología Cuantitativa. Estrategias y técnicas de investigación social.* Ed. Síntesis. Madrid.

Cortada de Kohan, Nuria. (1999). *Teorías Psicométricas y Construcción de Test.* Lugar Editorial. Buenos Aires.

Diccionario de la Lengua Española (2001). Real Academia Española. 22ª Edición. <http://www.rae.es/rae.html>.

Kaplan R. y D. Saccuzzo (2006). *Pruebas psicológicas.* Ed. Thomson.

Hirsch, A. (2005). Construcción de una escala de actitudes sobre ética profesional. *Revista Electrónica de Investigación Educativa*, 7 (1). Consultado el 14/6/2008, en: <http://redie.uabc.mx/vol7no1/contenido-hirsch.html>.

Pérez Ibarra, M.; Pérez Otero, N.; Méndez, S.; García, A.; Gálvez Díaz, M.; Quincoces, V.; Liberatori, H.; Fiorito, B.; Lasserre, C. (2007). "Construcción y Validación del cuestionario: Métrica de Calidad de Credibilidad e Interacción de Cursos de Teleformación". *Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología*. Número 2, Julio 2007, pp30-38. <http://teyet-revista.info.unlp.edu.ar/files/No2/TEYET2-art03.pdf>

Sampieri, R.; Fernández-Collado, C. y P. Baptista Lucio (2007). *Metodología de la Investigación.* Ed. Mc Graw Hill. México.

Vieytes, R. (2004). *Metodología de la investigación en organizaciones, mercado y sociedad. Epistemología y técnicas.* Ed. De las Ciencias. Buenos Aires.

Zegers, B.; Larrain, M.; Polaino-Lorente, A. et al. Validez y confiabilidad de la versión española de la escala de Cohesión y Adaptabilidad Familiar (CAF) de Olson, Russell & Sprenkle para el diagnóstico del funcionamiento familiar en la población chilena. *Revista chilena neuropsiquiátrica.* [online]. Enero 2003, Vol.41, Nro.1 [citado 10 Abril 2008], p.39-54.