

Análisis multivariante, conceptos y aplicaciones en Psicología Educativa y Psicometría

*Antonio Humberto Closas, Edgardo Alberto Arriola,
Cristina Isabel Kuc, Mariela Rosana Amarilla,
Ethel Carina Jovanovich*

Resumen

El objetivo de este trabajo radica, por una parte, en brindar un panorama introductorio sobre distintos conceptos de Análisis Multivariante (AM) y, por otra, proporcionar información acerca de diferentes aplicaciones realizadas en el área de la Psicología Educativa y de la Psicometría. La disponibilidad de programas informáticos, el progreso de la tecnología y el desarrollo de la teoría multivariante, han producido en las últimas décadas un notable crecimiento en la utilización de métodos y modelos estadísticos en todos los campos de la investigación científica. Las técnicas que conforman el AM permiten abordar distintos problemas de predicción, clasificación y segmentación, habitualmente presentes en las realidades de diversas áreas de conocimiento.

Palabras clave: análisis multivariante - modelización estadística - psicología educativa - psicometría.

Abstract

The aim of this work lies, on one hand, in showing an introductory view about different concepts of Multivariate Analysis (MA), and, on the other hand, in providing information about different applications made in the area of Educational Psychology, and Psychometrics. The availability of computer programs, the progress of technology, and the development of the multivariate theory have produced a remarkable growth in the use of methods and statistic models in all the fields of scientific investigation in the last decades. The techniques which make up the MA allow to approach different problems of prediction, classification and segmentation, which are usually present in many areas of knowledge.

Keywords: multivariate analysis - statistics modeling - educational psychology - psychometrics.

Introducción

En diferentes campos del conocimiento, el análisis numérico de la información frecuentemente se realiza por medio de distintos cálculos estadísticos. Actualmente, se encuentran disponibles métodos que otorgan nuevas posibilidades de tratamiento cuantitativo, que ciertamente no sería posible realizar con los procedimientos tradicionales uni y bivariantes. Estos métodos, integrados por una serie de técnicas de análisis de datos que forman parte

de la rama de la Estadística conocida como *análisis multivariante*, resultan de gran utilidad para llevar adelante estudios tanto de dependencia como de interdependencia entre variables. Podría decirse que el análisis multivariante se inició con la utilización de la regresión lineal por parte de Gauss en 1809 y, posteriormente, por otros estadísticos como Markov en 1900; en tanto que las técnicas más recientes tuvieron origen en los años treinta. En el presente, se encuentran disponibles diversos programas informáticos (R, SPSS, S-PLUS, AMOS, STATA, MPLUS, SAS, EQS, SAS CALIS, entre otros), los cuales incluyen procedimientos razonablemente sencillos para aplicar la mayoría de las técnicas del análisis multivariante de datos.

En las últimas décadas se ha producido un gran crecimiento en la utilización de estas técnicas en todos los campos de la investigación científica. Si bien podrían darse muchas razones para este uso creciente, las tres más importantes –de acuerdo con Dillon y Goldstein– posiblemente sean las siguientes:¹

- En la mayoría de las investigaciones científicas es necesario analizar relaciones simultáneas entre tres o más variables.
- La investigación científica es un proceso iterativo, primero es necesaria la formulación explícita de las hipótesis que después han de contrastarse mediante la recogida y el análisis de los datos. Probablemente, los resultados de estos análisis sugieran una modificación de las hipótesis, siendo normal que durante este proceso se añadan y eliminen continuamente variables. La complejidad de los fenómenos analizados hace que sean muchas las variables implicadas, de allí que las investigaciones sean necesariamente multivariantes.
- El desarrollo de computadoras personales con capacidad de almacenamiento y potencia de procesamiento suficiente, así como de programas estadísticos cada vez más fáciles de usar y con mayores prestaciones para el análisis de datos.

Definir el análisis multivariante no es una tarea sencilla; no obstante, se dan algunas opciones que permitirán clarificarlo razonablemente y acercarse a su significado.

¹ W. R. Dillon y M. Goldstein, *Multivariate Analysis. Methods and Applications* (New York: Wiley & Sons, 1984).

- Ciertos autores se inclinan por mostrarlo como una extensión del análisis bivariante.² Bajo esta perspectiva, el análisis multivariante sería el caso general y las técnicas univariantes o bivariantes serían los casos particulares de las técnicas multivariantes.
- De una manera algo más formal, Kachigan define el análisis multivariante como la rama del análisis estadístico que se centra en la investigación simultánea de dos o más características (variables) medidas en un conjunto de objetos (personas, cosas o entidades).³
- En tanto que, para Salvador Figueras puede ser definido como el conjunto de métodos estadísticos cuya finalidad es analizar simultáneamente grupos de datos multivariantes, en el sentido de que hay varias variables medidas para cada individuo u objeto de estudio.⁴ Para este autor, su razón de ser radica en un mejor entendimiento del fenómeno que se analiza, obteniendo información que los métodos estadísticos univariantes y bivariantes son incapaces de conseguir.

En virtud de lo que antecede, se podría decir que las técnicas multivariantes difieren de las uní y bivariantes en que dirigen su atención, no al análisis de la media y la varianza de una variable, o a la correlación entre dos variables, sino al análisis de las covarianzas o correlaciones que reflejan las relaciones entre tres o más variables.

Los *objetivos* que persigue el análisis multivariante, podrían sintetizarse del siguiente modo:

- Proporcionar métodos cuya finalidad es el estudio conjunto de datos multivariantes, lo que a través del análisis estadístico uní y bivalente no es posible lograr.
- Ayudar al investigador a tomar decisiones óptimas en el contexto en el que se encuentre teniendo en cuenta la información que aporta el conjunto de datos analizado.

Antes de finalizar con esta breve introducción, se considera conveniente señalar que, al trabajar con datos reales, en muchas ocasiones resulta que las

² B. G. Tabachnick y L. S. Fidell, *Using Multivariate Statistics* (New York: Harper Collins, 1996, 3ª ed.).

³ S. K. Kachigan, *Multivariate Statistical Analysis* (New York: Radius Press, 1991, 2ª ed.).

⁴ M. Salvador Figueras, "Introducción al Análisis Multivariante" (2000 [citado el 31 de octubre de 2011]), disponible en <http://ciberconta.unizar.es/LECCION/anamul/inicio.html>

hipótesis de normalidad y homocedasticidad multivariante no se verifican estrictamente (en general, para la aplicación de las técnicas multivariantes se requiere el cumplimiento de los supuestos paramétricos: normalidad, linealidad, independencia, homocedasticidad, etc.). Sin embargo, en la mayoría de estas situaciones dichos supuestos “casi” se cumplen; de todas formas, lo que realmente interesa, en estos casos, es conocer los efectos que pueden tener sobre los resultados las violaciones de estas hipótesis.

Existe clara evidencia, tanto matemática como empírica, de que las pruebas con una sola variable dependiente (análisis de la varianza, de regresión, de la covarianza, etc.) son altamente robustas bajo la violación de los principios de normalidad y homocedasticidad, excepto cuando las muestras son muy pequeñas y desiguales. En el análisis multivariante, sostiene Harris, la tendencia está en considerar que en muestras grandes ($N > 30$) sus métodos son lo suficientemente robustos como para ser insensibles a ligeras desviaciones de los supuestos estadísticos.⁵

Sin embargo, siguiendo a Closas, en algunas técnicas las exigencias podrían ser mayores; por ejemplo, en modelos de ecuaciones estructurales (MEE), el tamaño muestral que habitualmente se requiere suele ser superior a los 100 objetos. Como regla general, en estudios mediante MEE, deben obtenerse cinco veces más observaciones que el número de variables que van a ser analizadas, aunque el ratio óptimo (observaciones-variables) para la selección de la muestra es de diez a uno.⁶

Se concluye este primer apartado señalando que el propósito de este trabajo, como de alguna manera dejan ver las consideraciones que anteceden, radica simplemente en brindar un panorama inicial, de tipo descriptivo-informativo, sobre distintos conceptos del análisis multivariante. Asimismo, se exponen breves aspectos acerca de ciertos trabajos de investigación aplicada, realizados en el campo de la Psicología Educativa y de la Psicometría, que se estiman pertinentes presentar en razón de que ilustran la utilidad de algunas de las técnicas que integran los conjuntos que serán objeto de estudio.

⁵ R. J. Harris, *A Primer of Multivariate Statistics* (Orlando, FL: Academic Press, 1985, 2ª ed.).

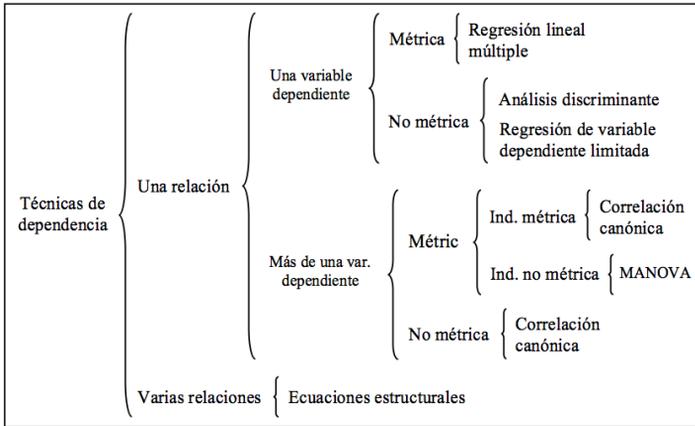
⁶ A. H. Closas, “Ecuaciones estructurales, una técnica estadística para formular y contrastar modelos de relaciones causa-efecto”, *Actualidad y Prospectiva* 1, n° 7 (Noviembre, 2011): 175-189.

Clasificación

Las técnicas multivariantes pueden ser clasificadas –de acuerdo con Uriel y Aldás– en dos grandes grupos:⁷

Técnicas explicativas o de dependencia

Estas técnicas investigan la existencia o ausencia de relaciones entre dos grupos de variables. En caso que estos grupos se encuentren clasificados en variables dependientes e independientes, el objetivo de las técnicas de dependencia será establecer si el conjunto de variables independientes afecta al conjunto de variables dependientes de manera conjunta o individualmente.



Fuente: Adaptado de Uriel y Aldás (2005).

Fig. 1. Técnicas explicativas o de dependencia.

En general, para saber qué técnica se debe aplicar es necesario conocer cómo están medidas las variables que participan en el estudio. El proceso de elección de algunas técnicas de dependencia, se realiza atendiendo a las respuestas que se brindan a las siguientes preguntas: (a) ¿las variables están medidas en escalas métricas o no métricas?, (b) ¿cuántas variables dependientes existen?, y (c) ¿cuántas relaciones se plantean entre las variables

⁷ E. Uriel y J. Aldás, *Análisis Multivariante Aplicado* (Madrid: Thomson, 2005).

dependientes e independientes? En este marco, las técnicas de análisis de dependencia son las que se indican en la Figura 1.

A continuación, se presentan algunos conceptos relativos a las técnicas multivariantes que se encuentran descritas en la Figura 1.

Regresión lineal múltiple

Es la técnica adecuada si en el análisis hay una variable dependiente cuyo valor depende de varias variables independientes métricas o no métricas. En diversas investigaciones esta técnica tiene un papel fundamental puesto que su aplicación permite, entre otras utilidades, observar de qué manera las variables independientes logran predecir la variable dependiente o criterio. En otras palabras, a partir del análisis de regresión será posible inferir acerca de la existencia o no de relaciones significativas entre las variables independientes y la variable criterio o respuesta, siempre dentro del marco científico establecido para la investigación en curso. El siguiente es un esquema simplificado relativo al análisis de regresión múltiple:

$$\underbrace{Y_1}_{\text{(métrica)}} \leftarrow \underbrace{(X_1, X_2, X_3, \dots, X_m)}_{\text{(métricas, no métricas)}}$$

A modo de ejemplo, se podría desear saber la existencia o no de relación entre el *nivel de ingresos*, medido en pesos, de una persona (variable dependiente métrica) y un grupo de variables entre las que se encuentran el *nivel educativo*, la *edad* (ambas independientes métricas) y el *género* (independiente no métrica).

Análisis discriminante

Esta técnica se utiliza para clasificar a distintos individuos en grupos (o poblaciones) alternativos a partir de los valores de un conjunto de variables sobre los individuos a los que se pretende clasificar. Cada individuo puede pertenecer a un solo grupo. La pertenencia a uno u otro grupo se introduce en el análisis mediante una variable categórica que toma tantos valores como grupos existan. En el análisis discriminante esta variable desempeña el papel de variable dependiente.

Regresión de variable dependiente limitada

Es la técnica que se utiliza en aquellos casos como los planteados para la regresión lineal múltiple; sin embargo, presenta la diferencia de que la variable dependiente es dicotómica (claro que en esta situación también podría usarse el análisis discriminante). Así, por ejemplo, puede suceder que las opciones de respuestas de la variable *nivel de ingresos*, antes citada, se encuentren codificadas como 1 = inferior a la media y 2 = superior a la media.

De lo expuesto surge que, más allá de coincidir en el objetivo de análisis, el modo en que está medida la variable dependiente condiciona la elección de la técnica. Esto es, si la variable dependiente es métrica se debe utilizar el análisis de regresión lineal; en caso contrario, se ha de optar por el análisis discriminante (si es categórica con varios grupos) o la regresión de variable dependiente limitada (si los grupos definidos a priori son exclusivamente dos).

Correlación canónica

Este análisis pretende determinar la existencia de asociación lineal entre un conjunto de variables independientes y otro conjunto de variables dependientes, de acuerdo con el esquema simplificado:

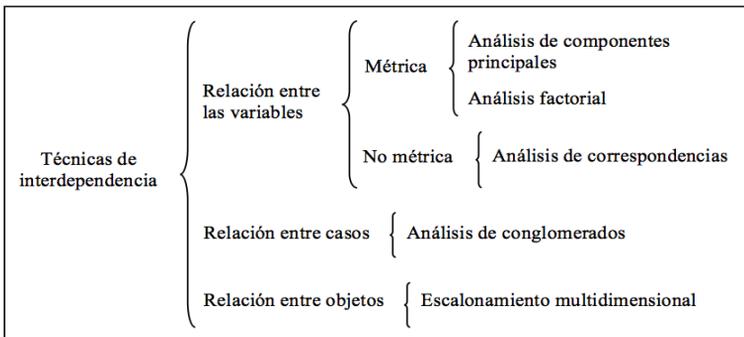
$$\underbrace{(Y_1, Y_2, Y_3, \dots, Y_n)}_{\text{(métricas, no métricas)}} \leftarrow \underbrace{(X_1, X_2, X_3, \dots, X_m)}_{\text{(métricas, no métricas)}}$$

A modo de ejemplo, un investigador puede querer establecer cómo influyen las variables antes mencionadas —*nivel educativo*, *edad* y *género*— no sólo sobre el *nivel de ingresos*, medido en pesos, sino también sobre el *nivel de satisfacción con el empleo actual*, medido mediante una escala tipo Likert en la que las opciones se valoran de 1 (totalmente insatisfecho) a 5 (totalmente satisfecho) puntos, por lo que ambas variables dependientes resultan métricas. En vista del ejemplo presentado, es fácil ver que la regresión lineal múltiple conforma un caso particular del análisis de correlación canónica cuando sólo se dispone de una variable dependiente.

El análisis factorial confirmatorio se utiliza en aquellos modelos en los que sólo están planteadas correlaciones entre variables latentes; en cambio, cuando además se postulan relaciones de causalidad en la que participan dos o más variables latentes, el análisis correspondiente se efectúa por modelos de estructuras de covarianza.

Técnicas descriptivas o de interdependencia

En caso de encontrarse ante una situación en la que sea imposible distinguir conceptualmente entre variables dependientes e independientes, y lo que interesa es determinar cómo y por qué las variables están correlacionadas entre sí; entonces, se debe recurrir a los métodos estadísticos de interdependencia, pues son los que abordan estas cuestiones. Las técnicas de análisis de interdependencia son las que se observan en la Figura 2.



Fuente: Adaptado de Uriel y Aldás (2005).

Fig. 2. Técnicas descriptivas o de interdependencia

Seguidamente, se brindan algunos aspectos de los métodos de interdependencia explicitados en la Figura 2 que, como se ha dicho, no distinguen entre variables dependientes e independientes y su objetivo reside en identificar qué variables están relacionadas, cómo lo están y por qué.

Análisis de componentes principales

Es una técnica de reducción de datos cuyo objetivo fundamental es construir combinaciones lineales de las variables originales que expliquen la mayor parte posible de la información recogida en estas variables. Cada combinación lineal se extrae de tal forma que están incorrelacionadas con las anteriores; además aporta, cada vez, menos información que la anterior.

Análisis factorial

También es una técnica de reducción de datos, como la anterior, pero a diferencia de ésta, el interés se centra en establecer qué variables latentes (factores) pueden estar causando la correlación entre las variables observadas. Puede verse como una técnica que pretende identificar grupos de variables de tal forma que las correlaciones entre las variables de cada grupo sean superiores a las correlaciones de las variables entre los grupos. En resumen, puede decirse que ambas técnicas se usan para analizar interrelaciones entre un número elevado de variables métricas explicando dichas interrelaciones en términos de un número menor de variables denominadas factores (si son inobservables) o componentes principales (si son observables).

Análisis de correspondencias

Esta técnica estadística pretende representar en un espacio multidimensional reducido la relación existente entre las categorías de dos variables no métricas. En el análisis de correspondencias, el mapa perceptual que se obtiene muestra las distancias entre los distintos niveles de dos variables, por lo que suele decirse que este análisis sirve para visualizar tablas de contingencia. Por ejemplo, si se desea saber si existe relación entre dos variables (*tipo de ocupación de un sujeto y universidad donde se graduó*), cada una de las cuales tiene 15 niveles. La tabla que se obtendría, de 15x15, impediría distinguir con facilidad lo que se desea conocer; sin embargo, el análisis de correspondencias permitiría mostrar un mapa, probablemente bidimensional, en el cual una universidad estará tanto más cerca de un tipo de ocupación, cuanto mayor sea la frecuencia de esa celda en la tabla de contingencia. Esta técnica, cuya teoría data de principios del siglo XX, pese a su innegable utilidad, no ha sido tan utilizada como otras en investigación en ciencias sociales. Este hecho curioso, posiblemente sea debido a que fue en los trabajos realizados

en Francia por Benzécri y sus colaboradores⁸ —un contexto no anglófono— donde ha tenido, en los años sesenta, su máximo desarrollo.

Análisis de conglomerados

El análisis de conglomerados, al que también se denomina comúnmente *análisis cluster*, es una técnica diseñada para clasificar distintas observaciones en grupos (diferente del análisis factorial, debido a que la finalidad de éste radica en agrupar variables), de tal forma que:

1. Cada grupo (conglomerado o cluster) sea homogéneo respecto a las variables utilizadas para caracterizarlos; es decir, que cada observación contenida en él sea parecida a todas las que estén incluidas en ese grupo.
2. Que los grupos sean lo más distintos posibles unos de otros respecto a las variables consideradas.

Escalamiento multidimensional

El análisis de escalamiento multidimensional (*MDS*, *multidimensional scaling*) es una técnica de reducción de datos, como lo es el análisis factorial o el análisis de componentes principales. El objetivo principal de esta técnica es representar N objetos en un espacio dimensional reducido (q dimensiones, siendo $q < N$), de tal forma que la distorsión causada por la reducción de la dimensionalidad sea la menor posible; es decir, que las distancias entre los objetos representados en el espacio q dimensional, sean lo más parecidas posible a las distancias en el espacio N dimensional.

A efectos de evaluar la proximidad de estas distancias, se utiliza un indicador que suele denominarse *stress* o *s-stress*. Un clásico ejemplo que se emplea para ilustrar la aplicación y objetivos de esta técnica, es la reconstrucción del mapa de un país con la posición aproximada de sus ciudades, a partir de una matriz conformada por las distancias en kilómetros entre aquellas más importantes de dicho país.

⁸ J. P. Benzécri, *L'Analyse des Données* (París: Dunod, 1973).

Proceso de aplicación de una técnica multivariante

Las técnicas multivariantes son, en su mayoría, herramientas muy poderosas que permiten al investigador extraer abundante información de los datos disponibles. Estas técnicas son, en sí mismas, relativamente complejas y requieren para su utilización un conocimiento profundo de sus fundamentos y condiciones de aplicabilidad. Sin embargo, el notable desarrollo de programas informáticos de manejo sencillo y el invariable avance de la tecnología están provocando que, en muchas áreas de conocimiento, y de manera especial en el campo de las ciencias sociales, las técnicas de análisis multivariante estén siendo ampliamente empleadas. El problema que de inmediato se presenta es que el uso indiscriminado de los paquetes estadísticos genera, muchas veces, que no se utilicen adecuadamente, puesto que el usuario con frecuencia carece del suficiente respaldo teórico que el escenario requiere.

En virtud de lo que precede, en este apartado se brinda una guía para la correcta aplicación de las técnicas multivariantes y el posterior arribo a conclusiones razonables. El desarrollo que a continuación se expone sigue el proceso planteado por Hair et al.⁹

1. Definición del problema que se está investigando, sus objetivos y decisión sobre la técnica multivariante que debe utilizarse: para ello se debe realizar una adecuada aproximación teórica al problema que se aborda; esto es, analizar conceptualmente el objeto de investigación, especificar los objetivos e identificar las relaciones fundamentales que se pretenden investigar. Sólo después de haber realizado lo anterior, se pueden seguir los esquemas representados en el apartado previo para seleccionar la técnica más conveniente.
2. Desarrollo del plan de análisis: una vez establecido el modelo conceptual, el énfasis debe centrarse en aplicar correctamente la técnica elegida, por lo que será necesario observar que la muestra posea un tamaño que permita su aplicación. También, debe asegurarse que los instrumentos que se usan en el procedimiento de recogida de datos (e.g., cuestionarios, escalas, tests) midan las variables con las escalas pertinentes (métricas y/o no métricas).
3. Análisis de las condiciones de aplicabilidad de la técnica elegida: tras haber recogido los datos, es necesario evaluar las hipótesis subyacentes

⁹ J. F. Hair, R. E. Anderson, R. L. Tatham y W. Black, *Multivariate Data Analysis* (Englewood Cliffs, NJ: Prentice-Hall, 1995, 4ª ed.).

a la técnica multivariante, puesto que si no se cumplen hace que carezca de sentido su aplicación. En las técnicas de dependencia, por ejemplo, suele ser necesario que los datos verifiquen hipótesis de normalidad, linealidad, independencia y homoscedasticidad. Además, en caso que los hubiera, debe decidirse qué hacer con los valores perdidos.

4. Estimación del modelo multivariante y ajuste global del mismo: se aplica la técnica multivariante elegida, luego se evalúa si el nivel de bondad del ajuste es adecuado. Si no es así, se deberá reespecificar el modelo, incorporando o eliminando variables. Se hace notar que no todas las técnicas tienen indicadores de bondad de ajuste.
5. Interpretación de los resultados: luego de que se logra un nivel de ajuste aceptable, se procede a interpretar el modelo. Es necesario prestar atención a los efectos de las variables individuales examinando sus coeficientes, cargas factoriales, utilidades, etc. La interpretación puede conducir a nuevas reespecificaciones del modelo.
6. Validación del modelo: previo a la aceptación de los resultados a los que se haya llegado, debe aplicarse una serie de técnicas de diagnóstico que aseguren que los resultados obtenidos pueden ser generalizados al conjunto de la población.

Aplicaciones

El propósito de este apartado es brindar, en forma breve, información sobre una serie de trabajos, producto de investigaciones en Psicología Educativa y Psicometría realizadas en los últimos seis años, en los que para lograr el objetivo propuesto fue necesario aplicar distintas técnicas de análisis multivariante, explicativas en algunos casos, descriptivas en otros y ambas en determinados estudios.

Investigaciones en Psicología Educativa

Relaciones entre aspectos cognitivo-motivacionales en estudiantes de Ingeniería y Ciencias Económicas

El objetivo de este trabajo consistió en analizar las tendencias motivacionales de las metas académicas (aprendizaje, refuerzo social, logro) que siguen los estudiantes universitarios y las relaciones que éstas presentan con otras variables —que contemplan aspectos cercanos a lo cognitivo y

motivacional—, como la *concepción de la inteligencia* (CI), la *capacidad percibida* (CP) y el *autoconcepto académico* (AA). Se utilizó una muestra compuesta por 437 alumnos de primer año matriculados, en el curso académico 2004, en la Facultad de Ingeniería y Ciencias Económico Sociales de la Universidad Nacional de San Luis, Argentina.

Los análisis de datos realizados, usando el paquete estadístico SPSS 12.0, se basaron inicialmente en el estudio de la consistencia interna del Cuestionario de Metas Académicas (CMA) —elaborado por Hayamizu y Weiner,¹⁰ empleado para medir los tipos de metas— mediante la evaluación de fiabilidad (*coeficiente alfa de Cronbach*) y la validez de constructo de la escala por el análisis factorial (*método de componentes principales, rotación varimax*). En segundo lugar, y en relación con el tema central de este estudio, se llevó a cabo un análisis correlacional multivariado (los resultados se muestran en la Tabla 1) con el fin de observar las relaciones que presentan los tipos de metas con las variables CI, CP y AA.¹¹

Tabla 1. Relaciones entre los tipos de metas y las variables CI, CP y AA

Variabes	M. de aprendizaje	M. de logro	M. de refuerzo social
CI	.188**	.014	-.064
CP	.148**	.011	-.017
AA	.235**	.106*	-.001

** $p < .01$; * $p < 0.05$ Fuente: Closas, Gatica y Pereyra (2006)

El análisis psicométrico del CMA muestra que el instrumento analizado es fiable (α de Cronbach para el total de la escala = 0.81) y posee validez de constructo (en conjunto las tres metas explican el 49% de la varianza total). Estos análisis permitieron conocer no sólo las características del CMA, sino también los tipos de metas académicas que persiguen los estudiantes en el contexto en el que se desarrolla la investigación.

Las relaciones entre metas académicas y las demás variables consideradas, parecen demostrar que aquellos estudiantes que sostienen una teoría de la inteligencia como algo que puede ser modificable, que se perciben

¹⁰ T. Hayamizu y B. Weiner, “A Test Dweck’s Model of Achievement Goals as Related to Perceptions of Ability”, *Journal of Experimental Education* 59, n° 3 (Spring, 1991): 226-234.

¹¹ Véase A. H. Closas, S. N. Gatica y S. N. Pereyra para más detalles acerca de las técnicas multivariantes citadas (“Relaciones entre aspectos cognitivo-motivacionales en estudiantes de Ingeniería y Ciencias Económicas”, *Revista de Ciencias de la Educación* 3, n° 207 [Julio-Octubre, 2006]: 271-292).

competentes y que además poseen una elevada valoración de su desempeño académico, tienden con mayor probabilidad a perseguir metas de aprendizaje. En cambio, los estudiantes que conciben la inteligencia como un aspecto invariable, disociado de la dedicación y el trabajo y que, al mismo tiempo, dudan de sus capacidades para llevar a cabo actividades académicas, demuestran predisposición a volcarse hacia metas de rendimiento (refuerzo social y logro).

Modelización estadística del rendimiento matemático con variables psicoeducativas en estudiantes universitarios

El propósito de esta investigación fue elaborar un modelo ajustado y representativo de las relaciones que se establecen entre ciertas variables de tipo personal y contextual que explique de qué manera las mismas influyen en el rendimiento en Matemáticas. La muestra estuvo compuesta por 441 alumnos (176 hombres y 265 mujeres) de la Facultad de Ciencias Económicas de la Universidad Nacional del Nordeste de Argentina, con una edad promedio de 20 años y desviación estándar de 2.94.

El desarrollo de este trabajo presenta la estructura clásica que tienen las investigaciones de este tipo. En el mismo fueron previstas dos partes, la primera corresponde al análisis teórico (permitió identificar las variables que, según la apreciación del autor de este trabajo, explican y modulan de manera significativa el rendimiento en Matemáticas) y la segunda se relaciona con el estudio empírico, en cuyo marco se plantea la discusión de resultados y las conclusiones.

En principio, se propone un *modelo teórico integral y jerárquico* mediante el cual se pretende explicar la varianza del problema que se investiga. En la etapa empírica, el proceso de análisis estadístico de datos revelados fases claramente diferenciadas: los análisis previos (descriptivos, de varianza, correlacionales, de regresión y factorial exploratorio) y los análisis confirmatorios (modelos de ecuaciones estructurales). Los primeros, situados en el dominio de la estadística descriptiva e inferencial, permitieron observar que, en general, el comportamiento de las variables, por un lado, armoniza aceptablemente con la propuesta inicial del modelo y, por otro, se encuentra en consonancia con los aportados por distintas investigaciones análogas. Los segundos, en cambio, pusieron en evidencia que si bien existe, casi siempre, validez de constructo de las variables latentes consideradas, se presentan

deficiencias estructurales, esto es, relaciones entre factores cuya inclusión en el modelo no podía ser empíricamente justificada (estudio analítico mediante el método de máxima verosimilitud). Además, los indicadores utilizados para juzgar globalmente el grado de ajuste resultaron en su mayoría inaceptables (diagnóstico de bondad de ajuste global).

Posteriormente, dada la falta de ajuste de la representación inicial fueron introducidas una serie de modificaciones, teóricamente coherentes y metodológicamente apropiadas, las cuales dieron lugar a un nuevo modelo que resultó ajustado a los datos en el que participan las variables: aspectos socio-familiares, estrategias de aprendizaje, autoconcepto académico, inteligencia general, características cognitivo-motivacionales y capacidades operativas. En esta última instancia, la estimación de las relaciones y la evaluación global del modelo fueron realizadas, como en la propuesta inicial, mediante el análisis de ecuaciones estructurales.

Desde el punto de vista analítico, tanto los coeficientes (componente de medida) como los parámetros (componente estructural), resultaron positivos y estadísticamente significativos (ver Figura 3). Esto es, por una parte, ha quedado probada la validez de constructo de los distintos factores considerados y, por otra, fue posible confirmar empíricamente que las relaciones postuladas entre las variables latentes presentan suficiente magnitud para ser incluidas en el modelo.

Respecto de los diversos criterios e índices utilizados [análisis de los residuos, convergencia en el proceso de estimación, estadístico χ^2 , razón χ^2/df , *Comparative Fit Index* (CFI), *Non-Normed Fit Index* (NNFI), *Root Mean Square Error of Approximation* (RMSEA), entre otros] para evaluar globalmente el ajuste del modelo, debemos señalar que los mismos sugieren su aceptación, dado que todos ellos verificaban razonablemente los valores referenciales que en cada caso se aplicaron [$\chi^2(50) = 52.57$, $p = .21$, $\chi^2/df = 1.05$, CFI = .99, NNFI = .98, RMSEA = .02].

En síntesis, tanto la cuantificación de las diversas relaciones propuestas, como los indicadores de bondad de ajuste, se mostraron adecuados y, por ende, pondrían de manifiesto que el modelo hipotetizado reproduce correctamente las asociaciones existentes en la matriz de correlaciones de datos empíricos. El modelo final propuesto sería, asimismo, un recurso válido para abordar con eficacia tareas de intervención educativa. Por último, deseamos

señalar que el análisis electrónico de los datos,¹² se realizó con ayuda de los programas SPSS (análisis previos) y EQS (análisis confirmatorios).

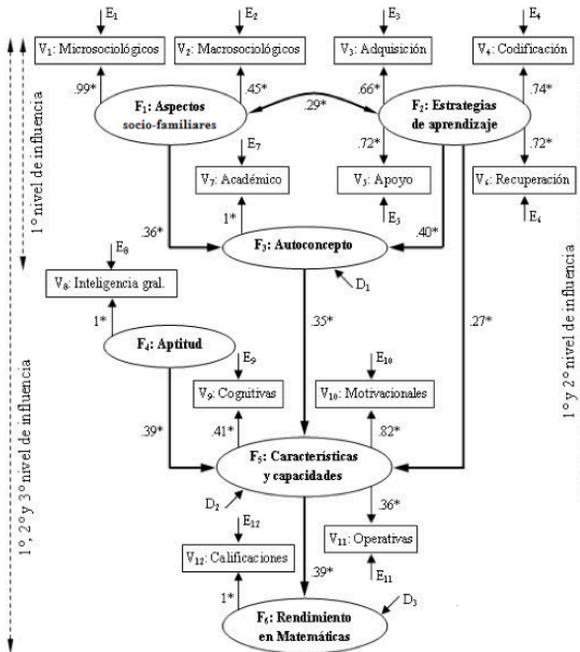


Fig. 3. Resultados estandarizados del modelo explicativo del rendimiento matemático.

Nota: Variables observadas: V1, V2; V3, V4, V5, V6; V7; V8; V9, V10, V11; V12. Variables latentes: F1, F2, F4 (independientes); F3, F5 (intermediarias); F6 (dependiente). *p < .05.

Fuente: Elaboración propia a partir de Closas (2009).

Investigaciones en Psicometría

Desarrollo y validación de un Cuestionario de Toma de Decisiones (Decision-Making Questionnaire)

El objetivo de este trabajo fue desarrollar y validar un cuestionario que sea de utilidad para examinar los factores que inciden en la “toma de

¹² Véase A. H. Closas para más detalles acerca de las distintas técnicas multivariantes utilizadas (“Modelización estadística del rendimiento matemático con variables psicoeducativas en estudiantes universitarios” [Tesis doctoral, Universidad Pública de Navarra, España, 2009]).

decisiones” (*DMQ*, *Decision-Making Questionnaire*). La investigación se llevó a cabo empleando dos muestras, una compuesta por 170 sujetos y otra por 425 participantes, de ambos sexos, todos ellos residentes en la Comunidad Foral de Navarra (España). Cada una de las muestras estaba dividida en tres intervalos de edades: jóvenes estudiantes (18-25 años), adultos (26-60 años), y personas de la tercera edad (61-75 años). Si bien el conjunto de análisis multivariantes utilizado en el estudio de los datos de esta investigación ha sido amplio: correlación, varianza, fiabilidad, factorial exploratorio y factorial confirmatorio, sólo se brindan algunas precisiones respecto de este último en razón de la relevancia que posee su aplicación en el artículo en cuestión.¹³

En efecto, para contrastar empíricamente la validez de constructo del DMQ se realizó un análisis factorial confirmatorio (método de máxima verosimilitud robusto, con el fin de corregir la ausencia de normalidad de los datos) sobre la matriz de varianzas y covarianzas de la segunda muestra, utilizando el programa EQS de Bentler.¹⁴ Cabe indicar que además de las opciones seleccionadas al momento de configurar las ecuaciones del modelo (formulación matemática), tanto de medida como estructural, fueron implementadas las siguientes decisiones. En primer lugar, se ha fijado a 1 la varianza correspondiente a cada uno de los factores comunes (*Task*, *Subject*, *Context*); en cambio, se han dejado libres las varianzas de los factores específicos (E_{Tj} , ..., E_{Wj}). En segundo lugar, se ha permitido que las covarianzas entre los factores comunes sean no nulas y, por el contrario, se han adoptado los supuestos de que los términos de error no correlacionan entre sí.

La evaluación se realizó por dos procedimientos. Por un lado, el estudio analítico, del que participan las diez variables observadas (*Uncertainty*, ..., *Work Pressure*) y las tres dimensiones del DMQ (*Task*, *Subject*, *Context*), con el objeto de estimar y contrastar las relaciones postuladas en el modelo; por otro lado, el análisis del grado de ajuste global del mismo, con el fin de comprobar en qué medida el modelo propuesto reproduce las relaciones existentes en la matriz de correlaciones de los datos empíricos.

Desde la perspectiva analítica, los resultados indican que la totalidad de las cargas factoriales estimadas, lo mismo que las covarianzas entre los

¹³ Véase M. L. Sanz de Acedo, M. T. Sanz de Acedo, M. Soria y A. H. Closas para más detalles relacionados con las restantes técnicas (“Development and Validation of a Decision-Making Questionnaire”, *British Journal of Guidance and Counselling* 37, n° 3 [August, 2009]: 357-373).

¹⁴ P. M. Bentler, *EQS Structural equations program manual* (Encino, CA: Multivariate Software, 2006).

factores comunes, son valores estadísticamente significativos y positivos. Efectivamente, en el marco del análisis de medición, fue posible comprobar que el DMQ posee una estructura compuesta por tres factores: *Task*, *Subject*, *Context*. Esto es, todos los coeficientes de regresión entre las variables observadas *Uncertainty*, *Time/Money*, *Consequences of Decision*, *Information & Goal* y el factor *Task*; de igual modo entre *Motivation*, *Self-Regulation*, *Cognition*, *Emotion* y el factor *Subject*; así como entre *Social Pressure*, *Work Pressure* y el factor *Context*, resultaron estadísticamente significativos ($p < .05$).

También las varianzas de los errores, excepto las correspondientes a E_{SP} y E_{WP} , resultaron significativas para un nivel $\alpha = .05$. A su vez, los resultados logrados señalan que las dimensiones *Task*, *Subject* y *Context* se encuentran correlacionadas en forma significativa y positiva ($p < .05$). Las distintas soluciones estandarizadas relativas al modelo propuesto, se muestran en la Figura 4.

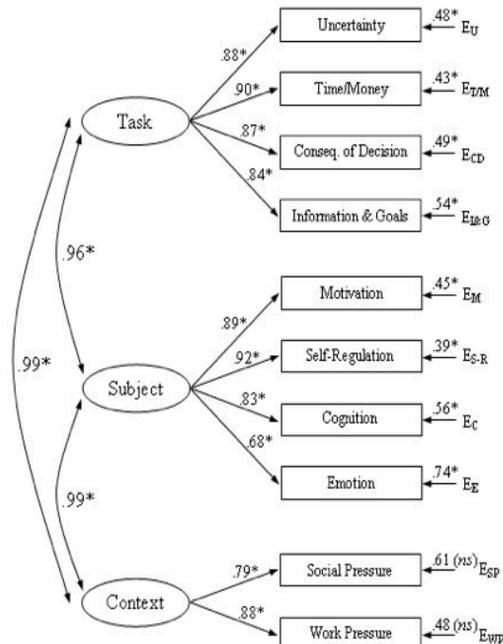


Fig. 4. Confirmatory factor analysis of the Decision-Making Questionnaire (DMQ) [Análisis factorial confirmatorio del Cuestionario de Toma de Decisiones (CTD)].

Note. ns = non significant effect, * $p < .05$.

Fuente: Sanz de Acedo et al. (2009)

A efectos de juzgar el ajuste global del modelo se ha asumido inicialmente, siguiendo a Schermelleh-Engel, Moosbrugger y Müller,¹⁵ una estrategia basada en cuatro índices: el estadístico χ^2 , junto con la razón entre χ^2 y los grados de libertad (χ^2/df), y otros indicadores descriptivos, que no dependen tanto del tamaño de la muestra, como CFI, NNFI y RMSEA, todos los cuales fueron citados en el trabajo presentado en segundo lugar en el apartado anterior. El tests chi-cuadrado resultó estadísticamente significativo, $\chi^2(32) = 58.05$, $p = .002$, lo que, en principio, podría hacer pensar en una posible inadecuación del modelo (si $\alpha = .05$); no obstante, el cociente χ^2/df , que suele acompañar al estadístico χ^2 debido a la especial sensibilidad de éste al tamaño muestral, siguiendo a Jöreskog y Sörbom,¹⁶ fue de 1.81. Los índices CFI y NNFI adoptaron valores .98 y .97, respectivamente; mientras que el límite inferior del intervalo confidencial para la RMSEA fue .02 y su estimación puntual, .04, indicativos todos ellos de un buen ajuste entre el modelo y los datos.

Por último, se añaden, a los índices de comparación estimados en primer término, los restantes estadísticos prácticos que proporciona EQS, entre los que se encuentran *Normed Fit Index* (NFI) = .95, *Incremental Fit Index* (IFI) = .98 y *Mc Donald's Fit Index* (MFI) = .97. Estos indicadores también dejan en evidencia que el modelo alcanzó bondad de ajuste, dado que superan el criterio de .90 recomendado por Bentler.¹⁷ En resumen, ha sido posible comprobar que la matriz de varianzas y covarianzas observada, y la predicha por el modelo especificado de acuerdo con las relaciones teóricas asumidas no son significativamente diferentes; por lo tanto, el modelo seleccionado se ajusta al modelo empírico.

Conceptualmente puede decirse que los análisis psicométricos revelaron una consistencia interna satisfactoria y estructura factorial de primer orden compuesta por diez escalas: incertidumbre, presión tiempo/dinero, información y metas, consecuencias de la decisión, motivación, autorregulación, cognición, emoción, presión social, y presión en el trabajo. A su vez, estas escalas han producido una estructura factorial de segundo orden, compuesta por: tarea, sujeto, y contexto. Los resultados de esta investigación ponen de

¹⁵ K. Schermelleh-Engel, H. Moosbrugger y H. Müller, "Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures", *Methods of Psychological Research Online* 8, nº 2, (August, 2003): 23-74.

¹⁶ K. G. Jöreskog y D. Sörbom, *Lisrel 8: Structural Equation Modeling with the SIMPLIS Command Language* (Hillsdale, NJ: Scientific Software International, 1993).

¹⁷ P. M. Bentler, *EQS Structural equations program manual* (Encino, CA: Multivariate Software, 2006).

manifiesto las formas en que los jóvenes, los adultos y las personas mayores se ven influidas por diversos factores al tomar decisiones. También sugieren que los hombres son más propensos que las mujeres a buscar y analizar información, mientras que las mujeres dan más valor a la disponibilidad de tiempo y dinero. El artículo concluye que el DMQ examina los factores que tienen mayor influencia en las decisiones profesionales y que podría guiar a los adolescentes en el proceso de decisión vocacional.

Validación de un cuestionario sobre estrategias de aprendizaje para estudiantes universitarios

En este trabajo se fijó como objetivo principal llevar a cabo la validación empírica de una versión abreviada de las Escalas de Estrategias de Aprendizaje, ACRA, elaboradas por Román y Gallego,¹⁸ con estudiantes de nivel universitario. El instrumento original es frecuentemente empleado en el ámbito hispanoparlante y está integrado por cuatro dimensiones que evalúan la utilización que habitualmente hacen los alumnos de las estrategias de adquisición, codificación, recuperación y apoyo (ACRA). Sin embargo, debido a su extensión y ámbito de aplicación, se planteó la posibilidad de seleccionar de esta prueba aquellos ítems que describan las técnicas mayormente utilizadas por los estudiantes, a efectos de lograr un instrumento que, con un formato más breve, resulte ajustado para ser aplicado en el nivel superior. La muestra seleccionada, utilizando los métodos estratificado, por conglomerados y aleatorio simple, estuvo compuesta por un total de 159 sujetos de primer año (55 mujeres y 104 hombres), pertenecientes a las tres carreras de Ingeniería que se imparten en la Facultad Regional Resistencia de la Universidad Tecnológica Nacional de Argentina, con una edad media de 19.91 años ($DT = 2.66$).

La investigación responde inicialmente a un diseño no experimental, explicativo y descriptivo mediante encuesta. También es una investigación de campo, de línea cuantitativa y de corte transversal. Los análisis de datos que se realizaron en la fase empírica pertenecen básicamente al dominio de la estadística descriptiva e inferencial: indicadores centrales y de dispersión, correlación ítem-total, fiabilidad, análisis correlacionales y análisis de regresión lineal.¹⁹ Estos cálculos permitieron, por un lado, conocer el comportamiento

¹⁸ J. M. Román y S. Gallego, *Escalas de Estrategias de Aprendizaje, ACRA* (Madrid: TEA, 1994).

¹⁹ Véase A. H. Closas, G. M. Sampallo, E. A. Arriola, C. I. Zening, M. R. Amarilla y E. C. Jovanovich para más detalles relacionados con la utilización de las técnicas mencionadas (“Validación de un cuestionario sobre estrategias de aprendizaje para estudiantes univer-

de cada una de las subescalas y de cada uno de los ítems que componen las categorías de la prueba aplicada; y, por otro, determinar la validez predictiva del instrumento, así como la ecuación de predicción que mejor describía la relación entre la variable criterio (Rendimiento académico) y las variables predictoras (dimensiones de las ACRA). En todos los casos, el procesamiento y análisis de los datos fue realizado con ayuda del programa informático SPSS 15.0.

En efecto, se comienza por señalar que los resultados relativos a los estudios descriptivos de los distintos ítems fueron muy similares a los informados por los autores del cuestionario simplificado,²⁰ lo que era de esperar en atención a que no se realizaron modificaciones de ningún tipo en las preguntas, ni en la estructura de la prueba. Los índices de consistencia interna hallados tanto para el total de la prueba (α de Cronbach = .93), como para cada una de las dimensiones fueron correctos en todos los casos (van de .70 a .91). Más aún, respecto de los indicadores α de Cronbach cuando se excluye el ítem, se puede señalar que se observan valores aceptables en la dimensión Estrategias de adquisición (oscilan entre .66 y .70), valores correctos en las subescalas Estrategias de codificación (van de .76 a .80) y Estrategias de recuperación (varían entre .74 y .77), mientras que los más altos se encuentran en la categoría Estrategias de apoyo (se hallan entre .89 y .91), todo lo cual sugiere que el cuestionario utilizado puede considerarse un instrumento con una fiabilidad aceptable.

En la Tabla 2 pueden apreciarse los valores de la *media*, la *desviación típica*, la *correlación ítem-total* y el *coeficiente alfa de Cronbach*, obtenidos por los individuos de la muestra en cada uno de los ítems del cuestionario utilizado. Los dos primeros estadísticos (*media* y *desviación típica*) son de mucha utilidad, puesto que el conocimiento de ambas medidas ayuda a comprender, entre otras cosas, la distribución de los datos de la muestra. El tercero de los estadísticos nombrados (*correlación ítem-total*), recoge el grado de relación que cada uno de los ítems posee con el total de la dimensión a la que pertenece, lo que puede considerarse un indicador de su grado de discriminación. La fiabilidad es una de las características fundamentales de un test, una de las formas de evaluarla es mediante el cuarto estadístico (*coeficiente alfa de Cronbach*) el cual

sitarios”, *Revista La UTN en el NEA: Investigación y Desarrollo en la Regional Resistencia* 1, n° 1 [Julio, 2012]: 75-86).

²⁰ J. De la Fuente y F. Justicia, “Escala de estrategias de aprendizaje ACRA-Abreviada para alumnos universitarios”, *Revista electrónica de investigación educativa y psicopedagógica* 2, n° 1 (Octubre, 2003): 139-158.

indica la precisión o estabilidad de los resultados; señala la cuantía en que las medidas de la prueba están libres de errores casuales.

Tabla 2. Estadísticos descriptivos y de correlación de los ítems medidos en pruebas

Dimensión	Ítem	Media	DT	Correlación Ítem-total	α de Cronbach sin el ítem
Estrategias de adquisición Media = 25.28 DT = 5.11 $\alpha = .70$	Ítem 1	2.66	.97	.22	.70
	Ítem 2	2.82	1.11	.46	.66
	Ítem 3	2.57	1.18	.31	.69
	Ítem 4	2.40	1.16	.45	.66
	Ítem 5	2.67	1.10	.42	.67
	Ítem 6	3.25	.93	.49	.66
	Ítem 7	3.33	.88	.44	.67
	Ítem 8	2.77	1.06	.29	.69
	Ítem 9	2.80	.97	.32	.69
Estrategias de codificación Media = 26.73 DT = 5.84 $\alpha = .80$	Ítem 1	2.82	.89	.40	.79
	Ítem 2	3.03	1.02	.45	.78
	Ítem 3	2.72	.96	.31	.80
	Ítem 4	3.23	.85	.44	.78
	Ítem 5	2.82	.96	.55	.77
	Ítem 6	2.30	1.02	.55	.77
	Ítem 7	2.46	1.09	.59	.76
	Ítem 8	2.23	1.13	.56	.77
	Ítem 9	2.67	.95	.36	.79
	Ítem 10	2.45	.93	.47	.78
Estrategias de recuperación Media = 36.90 DT = 6.61 $\alpha = .77$	Ítem 1	2.96	.88	.51	.74
	Ítem 2	2.89	.92	.45	.75
	Ítem 3	2.82	.90	.36	.76
	Ítem 4	2.69	1.00	.39	.76
	Ítem 5	2.61	1.03	.30	.77
	Ítem 6	3.15	.90	.41	.75
	Ítem 7	2.93	.94	.43	.75
	Ítem 8	3.01	.97	.36	.76
	Ítem 9	3.17	.84	.43	.75
	Ítem 10	2.69	.97	.43	.75
	Ítem 11	2.24	1.01	.30	.77
	Ítem 12	2.81	.94	.42	.75
	Ítem 13	2.73	.95	.38	.76

Tabla 2. Estadísticos descriptivos y de correlación de los ítems medidos en pruebas (cont.)

Dimensión	Ítem	Media	DT	Correlación Ítem-total	α de Cronbach sin el ítem
Estrategias de apoyo Media = 58.85 DT = 14.35 $\alpha = .91$	Ítem 1	2.68	1.00	.57	.90
	Ítem 2	2.78	.97	.60	.89
	Ítem 3	2.74	1.04	.56	.90
	Ítem 4	2.97	1.05	.63	.89
	Ítem 5	2.78	1.04	.51	.90
	Ítem 6	2.52	1.07	.51	.90
	Ítem 7	2.51	1.18	.41	.91
	Ítem 8	2.54	1.20	.41	.91
	Ítem 9	2.89	.97	.55	.90
	Ítem 10	2.62	1.01	.49	.90
	Ítem 11	3.09	1.13	.57	.89
	Ítem 12	2.82	1.07	.49	.90
	Ítem 13	2.51	1.17	.53	.90
	Ítem 14	2.74	1.30	.45	.90
	Ítem 15	3.15	1.16	.64	.89
	Ítem 16	3.05	1.18	.58	.89
	Ítem 17	3.18	1.05	.60	.89
	Ítem 18	2.86	1.19	.55	.90
	Ítem 19	3.04	1.06	.61	.89
	Ítem 20	3.08	1.16	.61	.89
	Ítem 21	2.55	1.21	.37	.91

Fuente: Closas et al. (2012).

A su vez, ha sido confirmada la presencia de relaciones estadísticamente significativas entre las distintas subescalas del cuestionario, así como entre éstas y la variable rendimiento académico en el área de Matemática (véase Tabla 3).

Tabla 3. Matriz de correlaciones

Variables	Estrat. de adquisición	Estrat. de codificación	Estrat. de recuperación	Estrat. de apoyo	Rendimiento académico
Estrat. de adquisición	1	.68**	.59**	.45**	.03
Estrat. de codi- ficación		1	.60**	.39**	.06
Estrat. de recu- peración			1	.53**	.04
Estrat. de apoyo				1	.16*
Rendimiento académico					1

* $p < .05$; ** $p < .01$ Fuente: Closas et al. (2012)

También, en vista de los resultados correlacionales, fue posible apreciar la validez predictiva, respecto del *Rendimiento matemático* (variable explicada), de la dimensión *Estrategias de apoyo* (variable explicativa) del cuestionario bajo estudio.

$$\text{Rendimiento en Matemática} = 2.77 + .06 \times \text{Escala de apoyo}$$

Otros datos de interés relativos a la estimación del modelo son: valor del coeficiente de correlación $R = .16$, decididamente bajo; más aún lo es el coeficiente de determinación corregido $\bar{R}^2 = .03$, por lo que la proporción de varianza total explicada por la regresión es irrelevante (claro que se debe ser prudente en interpretar esta apreciación puesto que en el modelo sólo interviene una de las cuatro dimensiones que conforman la prueba). Este resultado puede deberse, entre otras cosas, a que la propia técnica es excesivamente restrictiva para un problema de esta índole y a que el modelo ha sido estimado a partir de datos de corte transversal. Por otra parte, el valor del estadístico de Durbin-Watson ($d = 1.99$), próximo a 2, indica que los términos de error adyacentes no están correlacionados.

En definitiva, el desarrollo de la investigación ha permitido no sólo disponer en principio de un instrumento reducido que resulta apropiado para medir en la educación superior un concepto de probada influencia en el rendimiento de los estudiantes, sino que además rescata la importancia de que las investigaciones que se realicen en esta línea deberían contextualizar sus resultados, a efectos de contar con nuevos marcos de referencia que permitan generalizar la aplicación –con ajustes o adaptaciones, si fueran necesarias– en el nivel universitario de la prueba objeto de análisis.

Si bien, en los estudios que se acaban de describir, se emplearon algunas de las técnicas mencionadas en el segundo punto de este artículo, se considera necesario señalar que en investigaciones realizadas en distintas ocasiones y en diferentes ámbitos disciplinares (sobre todo en áreas de ciencias sociales) se han utilizado otros métodos, no indicados en los trabajos expuestos, tales como análisis discriminante (técnica explicativa), análisis de correspondencias y análisis de conglomerados (técnicas descriptivas). En todos estos casos, como en los citados en el presente desarrollo, la experiencia de haber participado ha sido sumamente positiva desde todo punto de vista, lo cual, muy probablemente, sea consecuencia de la relevancia atribuible a la integración de aspectos conceptuales y metodológicos del análisis multivariante y al procesamiento electrónico de los datos mediante la utilización de eficientes equipos y programas informáticos.

Discusión y conclusiones

En este trabajo se había propuesto, por una parte, brindar un panorama introductorio sobre distintos conceptos del análisis multivariante y, por otra, proporcionar información en forma breve acerca de algunas aplicaciones en el área de Psicología Educativa y Psicometría, las cuales fueron implementadas con el fin de lograr los objetivos planteados en cada una de las investigaciones presentadas y en las que, de diferente manera, se ha tenido la oportunidad de intervenir.

Pues bien, en atención tanto al desarrollo de los conceptos, como a la presentación de las aplicaciones realizadas del tema objeto de estudio, se considera que la meta perseguida, más allá de algunas debilidades que difícilmente puedan evitarse en este tipo de elaboraciones, ha sido razonablemente lograda.

El análisis multivariante es, en esencia, un conjunto de técnicas orientadas al estudio de varias variables de modo simultáneo, con el objeto de comprender mejor el fenómeno o evento que se analiza. Esta situación se presenta debido a que mediante los métodos disponibles es posible generar una explicación más rica y cercana a la realidad, lo que deriva, a su vez, en un mayor conocimiento del fenómeno y, por lo tanto, en un escenario apropiado para la selección óptima de la decisión o respuesta de interés.

En vista de ello, y de las diferentes apreciaciones realizadas en distintos momentos de este trabajo, se podría señalar que el desarrollo de la teoría multivariante—que se inició realmente en el siglo XX—, junto con sus aplicaciones—surgidas en la década del treinta—, así como el notable y sostenido progreso de la tecnología y de programas estadísticos, son componentes de un mismo conjunto que usados en forma complementada permiten llevar adelante numerosas investigaciones, explicar muchos fenómenos, alcanzar diversos objetivos y probar diversidad de hipótesis, todo lo cual de otra manera no hubiera resultado posible.

Evidentemente, las técnicas que conforman el análisis multivariante poseen en la actualidad una enorme importancia puesto que, además de la utilidad que tradicionalmente han tenido, forman parte del grupo de métodos, más o menos sofisticados, que se encuentran englobados en el concepto de minería de datos (*DM, Data Mining*). Este hecho, en razón de la vigencia y utilización de la DM en diferentes áreas, como Educación, Psicología, Biometría, Ecología, Economía, Marketing, etc., a efectos de abordar problemas de predicción,

clasificación y segmentación, entre otros, exige ciertamente de mayores comentarios respecto de su relevante riqueza y validez aplicativa.

Sin embargo, en el empleo de las técnicas para el análisis de datos multidimensionales, existen dos cuestiones que, se recuerda, deben observarse, puesto que habitualmente se presentan cuando se utiliza el modelado estadístico como metodología empírica de investigación.

La primera de ellas, quizás muy evidente, es que resulta esencial dar al conocimiento de la teoría del área de aplicación la máxima importancia, ya que de este saber surgirá la formulación de los objetivos y de las hipótesis sustantivas —en cierto modo, de las directrices de la investigación— que orientarán la selección del proceso de modelización subyacente.

La segunda cuestión tiene que ver con el hecho de que las últimas versiones de los paquetes informáticos presentan interfaces con el usuario de uso bastante coloquial, por lo que el trabajo con modelos estadísticos ha experimentado una simplificación importante que continuamente está generando una mayor popularidad de éstos. Sin embargo, esta cordial realidad trae consigo el inconveniente de que, en algunas ocasiones, la utilización de dichos asistentes se realiza de manera incorrecta, debido a que no siempre se cuenta con el conocimiento técnico necesario para implementar adecuadamente su aplicación.

En definitiva, y más allá de la prudencia que se debe tener sobre estos dos temas, la posibilidad de contar en estos tiempos con métodos y modelos multivariantes, unida a la gran ventaja que brindan el empleo de computadoras personales y de eficientes programas para el proceso de análisis de datos, son circunstancias sumamente significativas que, sin duda, han impulsado desde los años setenta, y en gran medida, el estudio de distintos fenómenos en diferentes disciplinas. Las producciones científicas, en este contexto metodológico y operativo, se habían debilitado en décadas anteriores, muy probablemente por no contar con la batería de opciones (procedimientos estadísticos y tecnología informática) adecuada para llevar adelante las investigaciones que la realidad educativa, económica y social requería en aquellos momentos.

Independientemente de sus debilidades, se espera que este primer acercamiento por nuestra parte, respecto de algunos conceptos y aplicaciones del análisis multivariante, pueda aportar algo más de luz a este tema que, se puede decir, posee en general una relativa dificultad de comprensión. Según nuestro modo de ver, la utilización de técnicas multidimensionales

en distintas acciones y disciplinas es sin duda favorable, en virtud de considerarlo un recurso metodológico muy conveniente y eficiente a efectos de dar explicaciones, tomar decisiones o proponer soluciones sobre diferentes problemas y situaciones, que el devenir de las circunstancias plantea diaria y continuamente, tanto en el ámbito científico como en la actividad profesional.

Antonio Humberto Closas
Facultad Regional Resistencia
Universidad Tecnológica Nacional
Resistencia, Chaco, Argentina
E-mail: hclosas@hotmail.com

Edgardo Alberto Arriola
Facultad Regional Resistencia
Universidad Tecnológica Nacional
Resistencia, Chaco, Argentina
E-mail: earriola2006@yahoo.com.ar

Cristina Isabel Kuc Zening
Facultad Regional Resistencia
Universidad Tecnológica Nacional
Resistencia, Chaco, Argentina
E-mail: criska440@hotmail.com

Mariela Rosana Amarilla
Facultad Regional Resistencia
Universidad Tecnológica Nacional
Resistencia, Chaco, Argentina
E-mail: prof.mariela@live.com.ar

Ethel Carina Jovanovich
Facultad Regional Resistencia
Universidad Tecnológica Nacional
Resistencia, Chaco, Argentina
E-mail: carijovanovich@yahoo.com.ar

Recibido: 30 de noviembre de 2011
Aceptado: 29 de julio de 2013