

Capítulo 8

VALIDEZ PREDICTIVA DE LOS CUESTIONARIOS MAPE-II Y EMA II.

Ignacio Montero García-Celay y Jesús Alonso Tapia

INTRODUCCION

En los capítulos 6 y 7 de este volumen se ha presentado el proceso mediante el cual se han construido los cuestionarios, qué sub-escalas se han identificado y con qué estrategia, así como la consistencia interna de las mismas. En este apartado aportaremos los datos que se han recogido para el establecimiento de la capacidad predictiva de las escalas que incluyen el MAPE-II y el EMA-II, tal como los hemos adaptado para la población de enseñanza media.

Antes de pasar a la justificación y exposición de los métodos y resultados obtenidos en este sentido, nos gustaría hacer una breve reflexión sobre el sentido general de los mismos dentro de los planteamientos que actualmente se postulan entre los teóricos de la evaluación.

Tradicionalmente se han venido manejando distintos conceptos de validez en relación con la construcción de instrumentos de medida en el ámbito de las ciencias sociales. Así, dentro del concepto general de validez, que hace referencia a la certeza obtenida de que un instrumento dado mida realmente aquello que dice medir, se ha venido distinguiendo, principalmente, entre validez de criterio, validez de contenido y validez de constructo. Los datos que vamos a exponer dentro de este apartado se enmarcarían dentro del primero de los conceptos. Sin embargo, si nos atenemos a las formulaciones que actualmente se están haciendo del tema (ver Gómez, 1986 para una revisión reciente), hay que considerar cualquier dato que se obtenga como dentro de un mismo continuo encaminado a probar la validez de constructo del instrumento, es decir, a interpretarlo a la luz de aquellas aportaciones teóricas que organizan y dan sentido a las puntuaciones obtenidas a través del instrumento de medida (Cronbach, 1980).

Desde esta perspectiva hay que entender el presente estudio sobre la capacidad predictiva de las escalas elaboradas dentro de los cuestionarios en un doble sentido:

Por un lado, trata de aportar luz al problema general de si estamos midiendo lo que queremos medir. Para ello se ponen a prueba hipótesis referidas a la relación existente entre este tipo de variables motivacionales y de estilos atributivos y el rendimiento en determinadas tareas; en nuestro caso en tareas de índole académico. Ahora bien, de ello no se desprende que las puntuaciones obtenidas en los cuestionarios **sólo** predigan el rendimiento académico, ni tampoco que **todos** los aspectos que se evalúan a través de las distintas subescalas aisladas sean buenos predictores del mismo. En este sentido, este estudio aporta luz al conocimiento que sobre el significado de las escalas podamos tener.

De otro lado, sin embargo, resulta interesante, a la vez, aportar datos sobre la utilidad predictiva en un ámbito concreto, sobre todo cuando dicho ámbito es -dado el número de materias utilizadas como criterio- lo suficientemente amplio, y adolece de instrumentos adecuados para la evaluación de los

aspectos motivacionales a los que venimos haciendo referencia.

JUSTIFICACION Y METODOS DEL ESTUDIO SOBRE VALIDEZ PREDICTIVA

Habiendo ya enmarcado el sentido general del estudio dentro de un continuo de datos sobre la validez de nuestros cuestionarios, vamos a continuar presentando los distintos métodos de estimación de la capacidad predictiva de las escalas utilizadas.

El objetivo último de este estudio es el de establecer de la forma más adecuada posible el grado en el que distintos conjuntos de variables predicen los criterios de rendimiento elegidos. Ahora bien, ¿qué predictores y qué criterios?

Predictores

Hemos establecido distintas agrupaciones de predictores manejando variables de motivación (cuestionario MAPE-II), estilos atributivos (cuestionario E.M.A.-II) y de inteligencia general (factor "g"). Las razones que justifican esta elección son varias. En primer lugar, obviamente, nos interesa saber la capacidad predictiva de las escalas motivacionales que hemos diseñado. En segundo lugar, también nos interesa estudiar su capacidad predictiva cuando se las combina con otro tipo de variables próximas a la motivación como son los estilos atributivos. Por último, hemos considerado importante ver la capacidad predictiva que añaden a la de variables de tipo aptitudinal como es el factor de inteligencia general.

Tanto el MAPE como el EMA tienen versiones anteriores de ahí su apéndice en números romanos. Sin embargo, para facilitar la exposición, en este capítulo y en el siguiente obviaremos tal apéndice dejando los números romanos para la denominación abreviada de las escalas de segundo orden de ambos cuestionarios.

Las combinaciones de predictores han sido:

A1 - Cada cuestionario (MAPE y EMA) de forma aislada empleando las puntuaciones en las escalas de primer orden. La razón fundamental para hacer tales agrupaciones de predictores es la de aportar datos a aquellas personas que utilicen alguno de los cuestionarios de forma aislada y atendiendo a los aspectos más específicos (escalas de primer orden), así como establecer de forma diferencial la capacidad predictiva de los componentes motivacional y atributivo.

A2 - Combinaciones de los tres tipos de variables (factor "g" y escalas de primer orden en el caso de los dos cuestionarios). En este caso, cabe justificar la obtención de la capacidad predictiva de todos los conjuntos de predictores (MAPE/EMA, MAPE/factor g, EMA/factor g y MAPE/EMA/factor g) desde dos planteamientos: a) La fundamentación que la teoría aporta sobre la interacción de los tres tipos de variables y b) el establecimiento del incremento que la utilización de las medidas motivacionales y de estilos atributivos añade a la capacidad predictiva de las medidas de inteligencia más frecuentemente utilizadas en el contexto académico.

A3 - Cada cuestionario (MAPE y EMA) de forma aislada utilizando las puntuaciones obtenidas

en las escalas de segundo orden, tanto las obtenidas de forma ponderada como las obtenidas por el método directo. En el caso del primer tipo de puntuaciones, su estudio como predictores se justifica por su mayor simplicidad de uso frente a las escalas de primer orden y por ser más fácilmente interpretables desde la teoría. En el caso de las segundas se justifica, además, al simplificar todavía más el procedimiento para su obtención. En cualquier caso ello queda condicionado a que no se aprecie una variación sustancial en la capacidad predictiva de los cuestionarios en función del tipo de puntuación utilizada.

A4 - Combinaciones del mismo tipo de las señaladas en el párrafo A2 pero utilizando las puntuaciones de segundo orden: ponderadas y directas. La justificación de tales agrupaciones es la misma que en el caso anterior, esto es, existen planteamientos teóricos que apoyan tales combinaciones de variables y además resulta interesante mostrar el incremento que supone la introducción de variables motivacionales y de estilos atributivos sobre la capacidad predictiva de las variables clásicas de inteligencia en el rendimiento académico. La utilización de estos conjuntos de predictores tendrá sentido en la medida en que no se den excesivas variaciones en cuanto a cantidad de varianza explicada con respecto a las predicciones realizadas con las puntuaciones de primer orden. Así mismo, se pretende establecer si la diferencia en el método de obtención de las escalas afecta también a su capacidad predictiva.

Criterios.

Como se puede deducir de lo hasta ahora expuesto, hemos utilizado como criterio el rendimiento académico de los sujetos, operativizado en la nota media final de todas las materias del curso en el que se les administraron las pruebas que se utilizan como predictores.

Dado que a partir de segundo curso, es decir en tercero y COU, los alumnos de bachillerato tienen la posibilidad de elegir cierto número de asignaturas en función de sus propios intereses hemos considerado útil añadir, para estos sujetos, dos criterios además de la nota media global: una nota media final de sus asignaturas optativas y otra de las asignaturas comunes obligatorias.

Como el grupo de sujetos de tercero y de COU era notablemente inferior al de la muestra total del estudio, y a fin de poder comparar de qué modo la capacidad predictiva de un conjunto de predictores varía de un criterio a otro hemos calculado dos predicciones distintas del criterio general de la nota media.

Además del estudio de validez predictiva sobre la nota media realizado con toda la muestra, hemos realizado otro semejante utilizando exclusivamente los sujetos de tercero y COU con los que se estudió la validez predictiva sobre los dos criterios parciales (nota media de optativas y nota media de comunes obligatorias), de tal modo que pudiéramos establecer comparaciones entre los tres criterios.

Una vez aclaradas las cuestiones acerca de cuáles han sido los predictores y criterios utilizados y por qué, quedan sin embargo algunas preguntas importantes que responder. ¿Cuál de los métodos y procedimientos existentes es el más adecuado para establecer la capacidad predictiva de las medidas

utilizadas? y ¿qué estabilidad tienen las relaciones obtenidas entre predictores y criterios?

Métodos.

Aunque la técnica que se recomienda en el caso de trabajar con un grupo de predictores y otro de criterios es la de la **correlación canónica** (San Martín y Espinosa, 1987), en nuestro caso hemos preferido la utilización de la técnica de análisis de **regresión múltiple** al considerar que nuestros criterios son en realidad distintas expresiones de un único criterio: el rendimiento académico. Hay que tener en cuenta que la correlación canónica no es más que una extensión de la regresión múltiple para el caso de la utilización de varios criterios distintos en el análisis.

Ahora bien, aún teniendo decidido el tipo de método a utilizar, existe el problema de que se pueden utilizar distintos procedimientos para el cálculo de la regresión múltiple. En nuestro caso, y para intentar controlar el efecto del procedimiento, hemos empleado dos de ellos -el directo y el iterativo o "stepwise" - para todos y cada uno de los análisis realizados.

Las razones que justifican el uso de esos dos métodos son las siguientes:

El método directo o estándar, tiene la ventaja de que minimiza los efectos del azar y permite ver la aportación de todos los predictores y su significación pero no permite conocer el incremento que la introducción de un nuevo predictor supone en términos de varianza explicada (R^2).

El método "stepwise" o iterativo produce sin embargo el efecto contrario, es más sensible a los efectos del azar en la composición de la muestra pero permite conocer la aportación de cada predictor al conjunto de la varianza explicada por la regresión.

Estabilidad de los pesos: Validación cruzada.

A la hora de utilizar las rectas de regresión obtenidas a través de una única muestra para predecir las puntuaciones en el criterio, conviene estudiar previamente la estabilidad de los valores que aparecen en las rectas mencionadas. Para ello se suele utilizar otra muestra de la que se tienen los valores obtenidos a través de las rectas que se obtuvieron con la primera muestra. A este procedimiento se le conoce como validación cruzada.

Existen casos -el nuestro es uno de ellos- en los que resulta muy difícil la obtención de otra muestra además de la utilizada en el estudio original. En tal caso lo que se recomienda es la selección aleatoria de una o varias submuestras de la muestra original y proceder con los datos de tales submuestras al cálculo de la validez cruzada siguiendo el procedimiento descrito anteriormente (San Martín y Espinosa, 1987; Tatsuska, 1976).

Así, hemos tomado dos submuestras, sujetos pares-impares, para la obtención de la validación cruzada y ello para cada uno de los grupos de predictores utilizados y los tres criterios.

Además de este procedimiento existen otros cuyos objetivos son establecer la cantidad de varianza que las rectas de regresión obtenidas predecirían en otras muestras. Dos son los métodos generalmente utilizados para ello: el método de Wherry y el de Stein (Tatsuska, 1976). La diferencia entre ambos métodos radica en que la proporción de varianza explicada en uno y otro caso se estima de

forma distinta. Con el primero de ellos, se hace a partir de la mejor combinación lineal de los predictores para la población globalmente entendida; con el método de Stein, sin embargo, esa óptima combinación de los predictores es realizada para la muestra normativa. La forma de optimizar la combinación lineal de las variables predictoras es maximizando las correlaciones al cuadrado. Por ello, la estimación que se realiza por el método de Wherry es siempre mayor que la que se lleva a cabo por el de Stein.

Método de Wherry: $R' = 1 - [(1-R)(N-1)/(N-p-1)]$

Método de Stein: $R'' = 1 - [(1-R)(N-1)(N-2)(N+1)/N(N-p-1)(N-p-2)]$

En ambas fórmulas, R hace referencia al valor de la regresión múltiple de la que parte la estimación, N es el tamaño de la muestra en la que se obtuvo el valor de R y p es el número de predictores utilizados en la ecuación de regresión.

Junto con los valores de la regresión obtenidos con cada uno de los grupos de predictores se presentarán los valores de las estimaciones obtenidas a partir de ambos métodos.

RESULTADOS

La exposición de los resultados obtenidos en los diversos estudios realizados para el cálculo de la validez predictiva de nuestro cuestionario está guiada por el mismo esquema que ha guiado la justificación de los mismos. Es decir, iremos mostrando los resultados alcanzados por los distintos grupos de predictores utilizados, después pasaremos a la comparación entre criterios para, finalmente, mostrar los datos obtenidos en cuanto a la estabilidad de las ecuaciones de regresión obtenidas. La comparación entre procedimientos de cálculo de la regresión -directo e iterativo- se realiza a la vez que se comparan los resultados que se han obtenido por los distintos grupos de predictores.

Primer grupo de predictores: Escalas de primer orden de cada cuestionario considerado independientemente.

En la tabla 8.1 se exponen los valores de R^2 que se han obtenido con los cuestionarios utilizados de forma aislada comparando los dos procedimientos: directo e iterativo. El criterio empleado es la nota media final del curso en todas las materias.

Dicha tabla pone de manifiesto tres cosas: En primer lugar que la diferencia entre procedimientos es prácticamente nula tanto por lo que se refiere a la cuantía del porcentaje de varianza explicada como por la significación de las variables que se incluyen en la ecuación de regresión, aunque en este segundo aspecto y en el caso del EMA se aprecia que el procedimiento iterativo incluye al primer factor si bien en último lugar y con una pequeñísima aportación a la varianza total (el 0,08 %). En segundo lugar, la capacidad predictiva del cuestionario de motivación es sensiblemente superior, casi un cuarenta por ciento más, a la del cuestionario de estilos atributivos. Posteriormente veremos que al no ser independientes las dos medidas el porcentaje de varianza que explican las dos juntas es mucho menor

que la suma de los porcentajes explicados por cada una de forma aislada. En tercer lugar, no todos los factores de los dos cuestionarios son buenos predictores del rendimiento académico. En el caso del MAPE quedan fuera de la ecuación los factores segundo y quinto, motivación intrínseca y ansiedad facilitadora respectivamente.

TABLA 8.1.: Resultados de la regresión múltiple sobre la nota media final del curso. Primer grupo de predictores. Métodos directo y "stepwise".				
Predictor	Método Directo		Método "Stepwise"	
M.A.P.E.	Coefi. Stand.	Prob.	Coefi. Stand.	Prob.
Escala 1	.212	<.0001	.215	<.0001
Escala 2	.021	=.5005	---	---
Escala 3	.128	<.0001	.121	<.0001
Escala 4	-.068	<.05	-.069	<.05
Escala 5	-.030	=.3093	---	---
Escala 6	-.214	<.0001	-.224	<.0001
Var. Expl.	R ² = .1848	<.0001	R ² = .1837	<.0001
E.M.A.	Coefi. Stand.	Prob.	Coefi. Stand.	Prob.
Escala 1	-.070	=.1023	-.093	<.05
Escala 2	-.141	<.0001	-.142	<.0001
Escala 3	-.042	=.2901	---	---
Escala 4	-.171	<.0001	.150	<.0001
Escala 5	.148	<.0001	-.181	<.0001
Escala 6	.024	=.4196	---	---
Escala 7	.163	<.0001	.165	<.0001
Var. Expl.	R ² = .1167	<.0001	R ² = .1152	<.0001

De nuevo se puede apreciar que las diferencias en cuanto al valor de R^2 obtenido mediante los dos procedimientos de cálculo de la regresión son prácticamente nulas (la máxima no alcanza el 0,8 % en valor absoluto) aunque en el caso del cuestionario EMA existen ligeras variaciones respecto de los factores que se incluyen en la ecuación dependiendo del procedimiento utilizado. En el caso del EMA quedan fuera, a su vez, el primero (según que procedimiento se utilice para el cálculo de la ecuación), el tercero y el sexto-atribución del éxito y del fracaso a factores aleatorios, atribución del éxito a causas no controlables y atribución del fracaso a la falta de habilidad. La explicación de este hecho hay que buscarla, en ambos casos, en dos direcciones: el significado del factor y la correlación existente entre los factores que han quedado fuera de la ecuación con alguno de los que se incluyen en la misma. Posteriormente volveremos a este punto más detenidamente.

Segundo grupo de predictores: Combinaciones de los cuestionarios (escalas de primer orden) y factor "g"

Los resultados correspondientes al cálculo de la regresión con los grupos de predictores que mencionamos en segundo lugar aparecen en la tabla 8.2. Al igual que en la tabla anterior se presentan las proporciones de varianza explicada que se han obtenido siguiendo los dos procedimientos de cálculo y se utiliza como criterio el rendimiento académico expresado en términos de nota media final de todas las asignaturas del curso.

El incremento de varianza que se explica utilizando las distintas combinaciones de variables con respecto a la que se obtiene con el uso de los cuestionarios de forma aislada es apreciable en todos los casos. Se aprecian también ligeras variaciones en el número y naturaleza de los factores que se incluyen en las ecuaciones de regresión al utilizar distintos grupos de predictores. Dichas variaciones afectan a aquellos factores que menor aportación hacían a la predicción en el caso anterior.

En el caso del cuestionario de motivación se aprecia una aportación importante sobre la cantidad de varianza explicada cuando se le añade a las aportaciones individuales del cuestionario de estilos atributivos EMA y al factor "g" (del 11,67% al 22,93 % y del 4,5 % al 21,99 %, respectivamente, en método directo). El incremento es mayor en el caso de la medida de inteligencia dado que la independencia entre los dos tipos de variables es más acusada que en el caso de los estilos atributivos. Cuando se añade a la combinación del EMA con el factor "g" su aportación sigue siendo importante (del 15,86 % al 25,31%), lo que supone casi un 10 % en términos absolutos y un 65 % en términos relativos. En todos los casos, cuando este cuestionario se agrupa en la ecuación con los otros dos, desaparece de la misma la cuarta escala (ansiedad inhibidora del rendimiento) que era la que menor incremento aportaba a la cantidad de varianza explicada por la regresión cuando se utilizaba el cuestionario de forma aislada en la predicción.

TABLA 8.2.: Resultados de la Regresión múltiple sobre la nota media final del curso. Segundo grupo de predictores. Métodos directo y "stepwise".

Predictor	Método Directo		Método "Stepwise"	
	Coefi. Stand.	Prob.	Coefi. Stand.	Prob.
Inteligenc.				
Factor "g"	.212	<.0001	.212	<.0001
MAPE + FG				
Factor "g"	.193	<.0001	.197	<.0001
Escala 1	.205	<.0001	.202	<.0001
Escala 2	.014	=.6528	---	---
Escala 3	.102	<.001	.092	<.01
Escala 4	-.035	=.2265	---	---
Escala 5	-.042	=.1550	---	---
Escala 6	-.225	<.0001	-.242	<.0001
Var. Expl.	R ² = .2199	<.0001	R ² = .2171	<.0001
EMA + FG				
Factor "g"	.208	<.0001	.206	<.0001
Escala 1	-.062	=.1392	-.085	<.05
Escala 2	-.126	<.001	-.127	<.001
Escala 3	-.044	=.2601	---	---
Escala 4	.169	<.0001	.170	<.0001
Escala 5	-.165	<.0001	-.176	<.0001
Escala 6	.036	=.2217	---	---
Escala 7	.131	<.0001	.135	<.0001
Var. Expl.	R ² = .1586	<.0001	R ² = .1509	<.0001
MAPE+EMA+FG				
Factor "g"	.183	<.0001	.186	<.0001
MAPE 1	.201	<.0001	.188	<.0001
MAPE 2	-.024	=.4488	---	---
MAPE 3	.080	<.05	.071	<.05
MAPE 4	-.018	=.5389	---	---
MAPE 5	-.017	=.5685	---	---
MAPE 6	-.186	<.0001	-.188	<.0001
EMA 1	-.080	<.05	---	---
EMA 2	-.115	<.001	-.133	<.001
EMA 3	.046	=.2184	---	---
EMA 4	.076	<.05	.083	<.05
EMA 5	-.101	<.001	-.096	<.01
EMA 6	-.003	=.9230	---	---
EMA 7	.086	<.01	.083	<.05
Var. Expl.	R ² = .2531	<.0001	R ² = .2488	<.0001

TABLA 8.2.: Resultados de la Regresión múltiple sobre la nota media final del curso. Segundo grupo de predictores. Métodos directo y "stepwise". (Continuación)

Predictor	Método Directo		Método "Stepwise"	
	Coefi. Stand.	Prob.	Coefi. Stand.	Prob.
MAPE + EMA				
MAPE 1	.206	<.0001	.196	<.0001
MAPE 2	-.018	=.6912	---	---
MAPE 3	.097	<.001	.101	<.001
MAPE 4	-.048	=.1055	---	---
MAPE 5	-.006	=.8412	---	---
MAPE 6	-.180	<.0001	-.189	<.0001
EMA 1	-.086	<.05	-.076	<.05
EMA 2	-.132	<.0001	-.131	<.0001
EMA 3	.048	=.2020	---	---
EMA 4	-.118	<.001	---	---
EMA 5	.064	<.05	-.095	<.01
EMA 6	-.009	=.7693	---	---
EMA 7	.100	<.001	.119	<.001
Var. Expl.	R ² = .2293	<.0001	R ² = .2229	<.0001

El cuestionario de estilos atributivos hace una aportación importante cuando se añade a la capacidad predictiva del factor "g" (del 4,5 % al 15,86 %) aunque esta aportación es apreciablemente menor a la que hacía el MAPE en el mismo caso. Con respecto a lo que aporta a la capacidad predictiva del MAPE y de la combinación MAPE/factor "g" (del 18,48 % al 22,93 % y del 21,99 % al 25,31 % respectivamente), es poco importante aunque apreciable. Se observan también algunas variaciones en cuanto a los factores que se incluyen en las ecuaciones de regresión, variaciones que ya se apreciaban en el caso de la comparación de los dos procedimientos de cálculo de la regresión utilizados cuando se usó el cuestionario como predictor de forma aislada. Parece que la pequeña cuantía que supone la introducción del factor primero de este cuestionario es sensible al tipo de procedimiento y a la introducción de nuevos predictores en la ecuación. De nuevo la explicación a este hecho hay que buscarla en relación a la existencia de correlaciones entre los factores que quedan dentro y fuera de la ecuación y al significado de los mismos.

Tercer grupo de predictores: Puntuaciones de segundo orden -ponderadas y directas- de cada cuestionario tomado de forma aislada

En la tabla 8.3 se presentan los resultados correspondientes al uso del tercer grupo de predictores. Como en los casos anteriores se incluyen los resultados alcanzados con los dos procedimientos de cálculo de la regresión y es utilizado el criterio general de la nota media final del curso de todas las materias. Además se incluyen los valores y ecuaciones de la regresión obtenidos por los dos tipos de escalas de segundo orden que se han venido proponiendo hasta ahora: aquellas que se obtienen ponderando la suma

TABLA 8.3: Resultados de la Regresión múltiple sobre la nota media final del curso. Tercer grupo de predictores. Métodos directo y "stepwise".

Predictor	Método Directo		Método "Stepwise"	
	Coefi. Stand.	Prob.	Coefi. Stand.	Prob.
MAPE (PSP.)				
Escala I	.358	<.0001	.358	<.0001
Escala II	.069	<.05	.069	<.05
Escala III	-.062	<.05	-.062	<.05
Var. Expl.	R ² = .1527	<.0001	R ² = .1527	<.0001
MAPE (P.F.)				
Factor I	.403	<.0001	.403	<.0001
Factor II	.072	<.05	.072	<.0001
Factor III	-.004	=.8796	---	---
Var. Expl.	R ² = .1680	<.0001	R ₂ = .1680	<.0001
EMA (PSP.)				
Escala I	-.195	<.0001	-.195	<.0001
Escala II	.232	<.0001	.232	<.0001
Escala III	-.132	<.0001	-.132	<.0001
Var. Expl.	R ² = .1042	<.0001	R ² = .1042	<.0001
EMA (P.F.)				
Factor I	-.184	<.0001	-.184	<.0001
Factor II	.197	<.0001	.197	<.0001
Factor III	-.112	<.001	-.112	<.0001
Var. Expl.	R ² = .0962	<.0001	R ² = .0962	<.0001

de las puntuaciones de primer orden y las que se obtienen en el caso de no realizar tal ponderación.

En este caso, más claramente aún que en los anteriores, se puede apreciar que las diferencias debidas al procedimiento utilizado en el cálculo de la regresión -directo o iterativo- es nula, tanto en lo que se refiere al porcentaje de varianza explicada como a los factores que entran en la ecuación. Donde si se aprecian diferencias es en la comparación de los resultados que obtienen los dos tipos distintos de escalas. Tales diferencias afectan tanto a la cuantía de varianza que explican como al número de factores que se incluyen en la ecuación.

En el caso del MAPE la ponderación supone una mejor capacidad predictiva y una menor disminución con respecto a la varianza que se explicaba con el uso de los factores de primer orden (se pasa de un 18,48 % a un 16,80 % con las ponderadas mientras que en el caso de las directas se desciende a un 15,27 %). Todo ello aún cuando en el caso de las escalas no ponderadas se incluye el factor tercero de segundo orden cosa que no

ocurre cuando las escalas se ponderan.

En general, la ponderación supone una mayor exactitud en la medición de los factores motivacionales y de su peso en la regresión frente al uso de las puntuaciones calculadas de forma directa ya que tiene en cuenta solamente la parte de varianza de cada escala de primer orden explicada por el factor correspondiente excluyendo el resto, cosa que no se hace en el caso de las puntuaciones directamente obtenidas.

En el caso del EMA ocurre exactamente lo contrario, si bien las diferencias entre los dos tipos de escalas son bastante menores (se pasa de un 11,67 % a un 9,62 % con las ponderadas frente a un 10,42 % con las directas). Las escalas no ponderadas predicen mejor el rendimiento y su uso supone un decremento menor con respecto a la capacidad predictiva del cuestionario cuando se utilizaban las escalas de primer orden. Los factores que se incluyen en la ecuación son los mismos en ambos casos.

Cuarto grupo de predictores: Combinaciones de cuestionarios (puntuaciones de segundo orden de los dos tipos) y factor "g"

Los resultados correspondientes a la utilización como predictores de los grupos mencionados en cuarto lugar se presentan en la tabla 8.4. Del mismo modo que en el caso anterior, se muestran la cantidad de varianza explicada y las variables que intervienen en la regresión para cada grupo de predictores incluyendo los dos procedimientos de cálculo de la misma así como las correspondientes a los dos tipos de escalas utilizadas dentro de cada cuestionario. El criterio empleado ha sido la nota media final de todas las materias del curso.

Se observa otra vez la ausencia de diferencias debidas al procedimiento de cálculo de la regresión. En cuanto al efecto que el tipo de escala produce, los resultados ponen de manifiesto la superioridad de las escalas ponderadas en todas aquellas combinaciones en las que interviene el MAPE, ocurriendo lo contrario - a semejanza de lo que se ponía de manifiesto cuando se utilizaban las escalas de primer orden- en aquella combinación (EMA/ factor g) en la que no interviene. En el caso del cuestionario de motivación, la diferencia entre tipos de escala es mayor cuando se combina con el factor "g", que en aquellos casos en los que participa el EMA en la combinación. Esto se explicaría al ser opuestos los efectos del tipo de escala en cada uno de los cuestionarios, lo que hace que al combinarlos se contrarresten.

En general, cabe señalar que el decremento de varianza que se produce al usar las diferentes combinaciones de escalas de segundo orden en lugar de las de primero no es distinto del que se producía al usar los predictores de segundo orden de cada cuestionario tomados independientemente en lugar de los de primer orden tomados, así mismo, de forma independiente.

Se aprecia de nuevo que la inclusión del MAPE junto al EMA o al factor "g" añade proporciones sustanciales a la capacidad predictiva de éstos tanto si se usan de forma aislada como si es su combinación la que se utiliza como predictor. Al lado de este hecho, también se aprecia que, aún cuando en el caso de su combinación con el EMA se mantienen los factores que aparecían en la ecuación cuando ésta se calculaba con el MAPE como único predictor, al combinarlo con el factor "g" desaparecen los factores segundo y tercero de las ecuaciones resultantes.

En el caso del EMA, el incremento que añade al porcentaje de varianza explicado por los otros dos predictores aislados -motivación e inteligencia- es parecido al que producía en el caso de utilizarse las escalas de primer orden. Es decir, incrementa de forma importante la capacidad predictiva del factor "g" aunque de forma inferior a la que lo hace el MAPE y añade poco a la capacidad predictiva del cuestionario de motivación aún cuando el incremento que aporta pueda considerarse apreciable.

TABLA 8.4: Resultados de la Regresión múltiple sobre la nota media final del curso. Cuarto grupo de predictores. Métodos directo y "stepwise".				
Predictor	Método Directo		Método "Stepwise"	
	Coefi. Stand.	Prob.	Coefi. Stand.	Prob.
Inteligencia				
Factor "g"	.212	<.0001	.212	<.0001
MAPE: PSP+FG				
Factor "g"	.190	<.0001	.201	<.0001
Escala I	.354	<.0001	.365	<.0001
Escala II	.038	=.1989	---	---
Escala III	-.031	=.2811	---	---
Var. Expl.	R ² = .1868	<.0001	R ² = .1846	<.0001
MAPE: PF+FG				
Factor "g"	.197	<.0001	.201	<.0001
Factor I	.389	<.0001	.389	<.0001
Factor II	.040	=.1654	---	---
Factor III	.021	=.4577	---	---
Var. Expl.	R ² = .2049	<.0001	R ² = .2029	<.0001
EMA: PSP+FG				
Factor "g"	.212	<.0001	.212	<.0001
Escala I	-.185	<.0001	-.185	<.0001
Escala II	.224	<.0001	.224	<.0001
Escala III	-.135	<.0001	-.135	<.0001
Var. Expl.	R ² = .1490	<.0001	R ² = .1490	<.0001
EMA: PF + FG				
Factor "g"	.212	<.0001	.212	<.0001
Factor I	-.176	<.0001	-.176	<.0001
Factor II	.189	<.0001	.189	<.0001
Factor III	-.108	<.001	-.108	<.001
Var. Expl.	R ² = .1408	<.0001	R ² = .1408	<.0001

TABLA 8.4: Resultados de la Regresión múltiple sobre la nota media final de curso. Cuarto grupo de predictores. Métodos directo y "stepwise". (Continuación).

Predictor	Método Directo		Método "Stepwise"	
	Coefi. Stand.	Prob.	Coefi. Stand.	Prob.
MAPE+EMA+FG	Coefi. Stand.	Prob.	Coefi. Stand.	Prob.
Factor "g"	.188	<.0001	.197	<.0001
MAPE Es I	.273	<.0001	.284	<.0001
MAPE Es II	.035	=.2636	---	---
MAPE Es III	-.029	=.3259	---	---
EMA Esc I	-.101	<.001	-.097	<.01
EMA Esc II	.141	<.0001	.148	<.0001
EMA Esc III	-.082	<.01	-.072	<.05
Var. Expl.	R ² = .2148	<.0001	R ² = .2131	<.0001
MAPE+EMA+FG	Coefi. Stand.	Prob.	Coefi. Stand.	Prob.
Factor "g"	.194	<.0001	.197	<.0001
MAPE Fc I	.320	<.0001	.332	<.0001
MAPE Fc II	.038	=.2005	---	---
MAPE Fc III	.019	=.5069	---	---
EMA Fac I	-.082	<.01	-.083	<.01
EMA Fac II	.102	<.001	.109	<.001
EMA Fac III	-.050	=.0972	---	---
Var. Expl.	R ² = .2221	<.0001	R ² = .2182	<.0001
MAPE + EMA	Coefi. Stand.	Prob.	Coefi. Stand.	Prob.
MAPE Es I	.278	<.0001	.278	<.0001
MAPE Es II	.061	<.05	.061	<.05
MAPE Es III	-.062	<.05	-.062	<.05
EMA Esc I	-.115	<.001	-.115	<.001
EMA Esc II	.142	<.0001	.142	<.0001
EMA Esc III	-.089	<.01	-.089	<.01
Var. Expl.	R ² = .1894	<.0001	R ² = .1894	<.0001
MAPE + EMA	Coefi. Stand.	Prob.	Coefi. Stand.	Prob.
MAPE Fc I	.333	<.0001	.333	<.0001
MAPE Fc II	.067	<.05	.068	<.05
MAPE Fc III	-.012	=.6717	---	---
EMA Fac I	-.089	<.01	-.092	<.01
EMA Fac II	.104	<.001	.103	<.001
EMA Fac III	-.063	<.05	-.061	<.05
Var. Expl.	R ² = .1936	<.0001	R ² = .1934	<.0001

Comparación entre criterios

Como ya se comentó al justificar este tipo de análisis, para poder realizar esta comparación se calcularon

las ecuaciones de regresión para los tres criterios únicamente con aquellos sujetos de la muestra que tenían la posibilidad de elegir entre grupos de asignaturas. Por ello la muestra utilizada en esta comparación es sensiblemente menor que en el caso de los análisis anteriores. Concretamente, han sido 353 sujetos, pertenecientes a tercero de BUP y COU, los que se han incluido en este estudio de comparación entre criterios.

En la tabla 8.5 se presentan los valores de R^2 y los factores que se incluyen en las ecuaciones de regresión correspondientes a los tres criterios utilizados: nota media global del curso, nota media de las asignaturas comunes y nota media de las asignaturas optativas. En este caso se presentan juntos todos los grupos de predictores que incluyen el cuestionario de motivación (MAPE) y no se incluyen comparaciones entre procedimientos de cálculo de la regresión ya que las realizadas con toda la muestra ponían de manifiesto la ausencia de diferencias apreciables entre los mismos. El procedimiento utilizado ha sido el directo.

TABLA 8.5.: Porcentaje de varianza explicada de cada uno de los criterios utilizados (Media global, comunes y optativas) por cada grupo de predictores en las que interviene el MAPE.

Predictores	Media global	Comunes	Optativas
Factor "g"		1.60 %	0.50 %
MAPE Esc. 1 Or.	15.61 %	13.63 %	16.58 %
MAPE Esc. 2 Or.	10.41 %	10.54 %	12.64 %
MAPE Fac. 2 Or.	13.25 %	12.49 %	15.20 %
MAPE E1Or. + FG	16.88 %	17.29 %	17.61 %
MAPE E2Or. + FG	11.26 %	13.60 %	13.35 %
MAPE F2Or. + FG	14.37 %	15.99 %	16.07 %
MAPE+EMA E1Or.	21.32 %	19.67 %	22.44 %
MAPE+EMA E2Or.	15.42 %	14.68 %	16.36 %
MAPE+EMA F2Or.	17.03 %	16.16 %	18.41 %
M+E (E1Or) + FG	22.05 %	22.37 %	23.18 %
M+E (E2Or) + FG	16.02 %	18.20 %	18.51 %
M+E (F2Or) + FG	17.83 %	19.81 %	20.01 %

Lo primero que cabe destacar es el hecho de que la utilización de una muestra más pequeña afecta de forma apreciable a la capacidad predictiva de los distintos grupos de predictores. En este caso la cantidad de varianza explicada tiende a ser menor que la que se explicaba mediante las ecuaciones obtenidas con todos los sujetos de la muestra. Esto puede interpretarse tanto en términos de estabilidad de las ecuaciones como desde un punto de vista evolutivo ya que los componentes de esta muestra pertenecen al segmento de mayor edad de todo el grupo de sujetos que han intervenido en el estudio anterior. La primera posibilidad se analiza de forma más clara en el próximo apartado. La segunda necesitaría de una investigación más sistemática para poder ser tenida en consideración. Recuérdese, en cualquier caso, que tampoco tenemos forma de saber si las diferencias que se observan son debidas al azar o tienen una significación estadística.

No vamos a comentar las diferencias que aparecen entre distintos grupos de predictores y distintos grupos

de escalas ya que lo hicimos anteriormente con los resultados de toda la muestra. Por lo que se refiere a las diferencias que aparecen entre los distintos criterios, dejando aparte el problema del establecimiento de la significación estadística de las mismas, cabe resaltar, fundamentalmente, dos hechos:

1.- La cantidad de varianza explicada es siempre mayor en el caso del criterio de la nota media final de las asignaturas optativas que el de la nota media final de las asignaturas comunes independientemente del tipo de predictores utilizados y del tipo de obtención de las puntuaciones para dichos predictores. Estos resultados parecen lógicos si se tiene en cuenta la naturaleza de ambos criterios. Cabe suponer que los sujetos estén más motivados en relación con las materias que eligen espontáneamente que en relación con las que deben cursar de forma obligatoria. El mismo razonamiento explicaría las diferencias que se observan al comparar la media en las optativas con la media global.

2.- La inclusión del factor "g" en las rectas de regresión favorece la capacidad predictiva más cuando el criterio es la nota media de las comunes que cuando es la de las optativas. Esto implica que esta variable de tipo aptitudinal influye más en la regresión cuando las asignaturas son obligatorias que cuando son elegidas por los alumnos.

Todo esto vendría a dar cierto soporte a nuestra suposición de que puede haber diferencias motivacionales importantes sobre el rendimiento mediadas por el hecho de la posibilidad de elegir las asignaturas a cursar. Aunque este es un hecho que se da por supuesto dentro del sistema educativo, nuestros datos vendrían a dar soporte de tipo empírico a tal creencia.

Estabilidad de las predicciones

Si se recuerda, proponíamos dos procedimientos para el estudio de la estabilidad de las ecuaciones elaboradas. Por un lado pensábamos utilizar las fórmulas diseñadas por Wherry y Stein (ver Tatsuoka, 1976) para estimar la cantidad de varianza que las ecuaciones predecirían en otras muestras del mismo tamaño y por otro planteábamos un estudio comparativo entre dos submuestras (sujetos pares e impares).

Por lo que se refiere al primero de los procedimientos que hemos utilizado, en la tabla 8.6 se muestran los resultados que se obtuvieron en toda la muestra con los que arroja la aplicación de los métodos de Wherry y Stein. No se han incluido las combinaciones de predictores que incluían el factor "g" y se mantienen las del EMA para poder tener un punto de comparación ya que las escalas de este cuestionario se muestran más estables que las del MAPE, que son las que fundamentalmente nos interesan en este trabajo.

Los resultados manifiestan que ambos métodos predicen una explicación de la varianza menor que en el caso de la muestra del estudio. Esta diferencia es mayor en el caso del método de Stein, tal como predecíamos al explicar la naturaleza de los dos métodos. También se observa las diferencias no son muy grandes (entorno a un uno por ciento en el caso de la mayor discrepancia) y que afectan más al MAPE que al EMA.

TABLA 8.6.: Estabilidad de los coeficientes de validez. Métodos de Wherry y Stein (Proporción de varianza explicada).			
Predictores	R^2 Obtenida	R^2' (Wherry)	R^2'' (Stein)

MAPE Esc. 1 Or.	.1848	.1800	.1750
MAPE Esc. 2 Or.	.1527	.1500	.1470
MAPE Fac. 2 Or.	.1680	.1660	.1620
EMA Esc. 1 Or.	.1167	.1110	.1040
EMA Esc. 2 Or.	.1042	.1020	.0982
EMA Fac. 2 Or.	.0962	.0936	.0902
MAPE+EMA E1Or.	.2293	.2200	.2090
MAPE+EMA E2Or.	.1894	.1850	.1790
MAPE+EMA F2Or.	.1936	.1890	.1840

Esta mejor estabilidad de las ecuaciones para el caso del cuestionario de estilos atributivos se pone también de manifiesto en la comparación de la cantidad de varianza que predicen las ecuaciones cuando se utilizan con las dos submuestras, de 547 sujetos impares y 546 sujetos pares, en las que hemos dividido a los sujetos de nuestra investigación (ver tabla 8.7). En el caso del MAPE, las oscilaciones existentes entre la cantidad de varianza explicada en toda la muestra y cada una de las dos submuestras llega a ser de tres puntos, estableciéndose un intervalo de oscilación de seis puntos entre ambas submuestras.

TABLA 8.7.: Estabilidad de los coeficientes de validez. Comparación entre submuestras (sujetos pares e impares) y muestra total. (R^2).			
Predictores	R^2 M.Total	R^2 Suj Pares	R^2 Suj Impares
MAPE Esc. 1 Or.	.1848	.1541	.2128
MAPE Esc. 2 Or.	.1527	.1235	.1835
MAPE Fac. 2 Or.	.1680	.1369	.1990
EMA Esc. 1 Or.	.1167	.0977	.1235
EMA Esc. 2 Or.	.1042	.1035	.1099
EMA Fac. 2 Or.	.0962	.0967	.0982
MAPE+EMA E1Or.	.2293	.2034	.2242
MAPE+EMA E2Or.	.1894	.1624	.1962
MAPE+EMA F2Or.	.1936	.1616	.2042

Cuando el MAPE se combina con el EMA en la ecuación la estabilidad mejora, probablemente debido a la mejor estabilidad que mostraba el cuestionario de estilos atributivos cuando era utilizado como único predictor.

En resumen, parece que la estabilidad de las ecuaciones es aceptable cuando se estima su capacidad predictiva en otras muestras del mismo tamaño, pero disminuye apreciablemente cuando el tamaño de la muestra se reduce a la mitad. Nuevamente insistir en el hecho de que no hemos encontrado modo de estudiar la significación estadística de las oscilaciones observadas.

CONCLUSIONES

La capacidad predictiva de los cuestionarios sobre el rendimiento académico ha resultado ser apreciable.

La cantidad de varianza explicada es mayor en el caso de utilizar las puntuaciones en las seis escalas como predictores que cuando se utilizan las de orden superior. En este caso resulta ser mejor la predicción cuando las puntuaciones se ponderan mediante los coeficientes factoriales que cuando se suman, sin más, las puntuaciones de las escalas de primer orden.

Cuando se combina el MAPE con otro tipo de predictores, como son los estilos atributivos o el factor de inteligencia general, el cuestionario de motivación aporta un porcentaje sustancial a la predicción tanto si la combinación es con una o las dos variables a la vez.

Sin embargo parece claro que no todas las escalas son buenos predictores del rendimiento académico. Dentro de las seis de primer orden del MAPE-II, la escala 2 -motivación intrínseca- y la 5 -ansiedad facilitadora del rendimiento- no hacen aportaciones significativas al porcentaje de la varianza del criterio que explica el cuestionario en su conjunto. En el caso de la primera de las dos mencionadas, este resultado es explicable debido a que no hay por qué esperar que un sujeto, que señala que cuando algo le gusta o interesa no hace falta que le den recompensas externas para trabajar en ello, esté, precisamente, interesado en las tareas escolares. En cuanto al caso de la escala de ansiedad facilitadora del rendimiento el resultado comentado vendría a poner en cuestión tal denominación. Si la mantenemos es debido a que su exclusión de la ecuación de regresión puede ser debida a su correlación con algún otro de los predictores y a que, en cualquier caso, los sujetos que puntúan alto en dicha escala poseen al menos la creencia de que cierto nivel de nerviosismo en situaciones difíciles les facilita su rendimiento en las mismas. Hay que señalar, además, que la escala de orden superior en la que se incluye la escala 5, predice significativamente el rendimiento.

La que no lo hace, a este nivel, es la escala III de miedo al fracaso cuando se pondera la puntuación mediante los coeficientes factoriales. Este hecho es más sorprendente si se tiene en cuenta que tal ponderación supone que la escala no difiera mucho de la 4 de primer orden -ansiedad inhibitoria del rendimiento- que sí hacía una aportación significativa a la predicción del rendimiento. Como ya hemos señalado, la utilización de un nombre distinto en cada caso es debido a la necesidad de distinguir su uso en el contexto de las escalas de primer orden o en el de las de segundo, pero consideramos que el significado de la puntuación es fundamentalmente el mismo. Como se propone una ecuación predictora para las escalas de los dos niveles, esta variable entrará en la predicción cuando ésta se realice con las escalas de primer orden pero no lo hará cuando se lleve a cabo con las segundas.

El cuestionario MAPE-II también se ha mostrado sensible en las predicciones del rendimiento académico cuando el criterio utilizado tiene en cuenta los distintos tipos de asignaturas -optativas y comunes-, poniendo de manifiesto que las variables motivacionales que evalúa predicen mejor el rendimiento en las primeras que en las segundas.

Finalmente, por lo que a la estabilidad de las predicciones se refiere, los resultados han mostrado que ésta es buena cuando se trata de hacer extrapolaciones para muestras del mismo tamaño, pero la capacidad predictiva sufre apreciables variaciones cuando se reduce el tamaño de la muestra a la mitad.