

# VALIDEZ INCREMENTAL: UN ESTUDIO APLICADO CON DIVERSAS FUENTES INFORMANTES Y MEDIDAS<sup>1</sup>

## INCREMENTAL VALIDITY: AN APPLIED STUDY USING DIFFERENT INFORMANTS AND MEASUREMENTS

CARRASCO<sup>2</sup>, M. A.; HOLGADO<sup>3</sup>, F. P.; DEL BARRIO<sup>2</sup>, M. V. Y BARBERO<sup>3</sup>, I.

<sup>2</sup>Dpto. Personalidad, Evaluación y Tratamiento Psicológico

<sup>3</sup>Dpto. Metodología de las Ciencias del Comportamiento

Facultad de Psicología

UNED

### Resumen

El estudio de la validez incremental es un importante tópico dentro de la evaluación psicológica pero aún está poco estudiado desde la psicometría. En el presente trabajo, se estudia la validez incremental de la agresión infantil informada por distintas fuentes (niño, compañeros y profesor) sobre la hostilidad y comunicación maternas. La muestra estudiada estuvo constituida por 525 sujetos (41,3 % varones) de edades comprendidas entre 9 y 14 años ( $M = 12,72$  años;  $DT = 1,57$ ) evaluados a lo largo de tres años consecutivos. Los resultados muestran el claro predominio de la fuente *autoinformada* frente a la *heteroinformada* en la predicción de la conducta materna, longitudinalmente estimada. Se discute el efecto artefactual de la fuente y los métodos psicométricos alternativos para una mayor precisión en el estudio de este tipo de validez.

### Palabras Clave

Validez incremental, regresión en orden jerárquico, agresividad, estilos de crianza.

### Abstract

The incremental validity study is an important psychological assessment topic that has been not study very much from psychometrics. In this study, we analyse the children incremental validity reported by different informants (children, classmates, and teachers). The sample comprised 525 schoolchildren (41.3 % boys) in a longitudinal study, ranging at age from 9 to 14 years ( $M = 12.72$  mean;  $SD = 1.57$ ) across the three consecutive years of measurement. In order to predict the maternal behaviour, self-reported source shows the higher incremental validity over hetero-reported source. We discuss alternative methods to study the incremental validity.

### Key Words

Incremental validity, hierarchical regression, aggressive, rearing styles, parenting behaviour

<sup>1</sup> Esta investigación ha sido subvencionada por el Ministerio de Ciencia y Tecnología dentro de la convocatoria I+D+I, 2001-2004 mediante el proyecto BSO2001 3042I.

## Introducción

La evaluación multimétodo y multifuente es una consigna aceptada por la mayoría de los evaluadores en psicología como parte de un proceder adecuado y útil, dado que, raras veces, una medida exclusiva o aislada es suficientemente comprensiva para proporcionar una evaluación completa (De los Reyes y Kazdin, 2004; Del Barrio, 2002; Fernández-Ballesteros, 1994; Garb, 2003; Meyer y Archer, 2001; Silva, 1995). Sin embargo, y aludiendo a la afirmación de McFall en relación con la evaluación multimétodo, *más no es siempre mejor (more is not always better)* (McFall, 2005, p. 316), la bondad de la simple combinación de medidas no debe ser asumida por defecto, debe ser demostrada empíricamente para cada objetivo de evaluación. Por tanto, el uso añadido de diferentes instrumentos, extensible al uso de diferentes fuentes, ha de ser contemplado, no sólo por su coste y grado de especificidad, sino también desde la perspectiva de la validez incremental, es decir, del grado en el que la inclusión de medidas e informantes a lo largo del proceso de evaluación mejore consistentemente la toma de decisiones y el poder predictivo (Garb, 2003; Hunsley, 2003; Hunsley y Mash, 2005). Parece razonable afirmar que, tanto el número como la combinación adecuada de instrumentos y fuentes informantes para la evaluación de un contenido en particular, llevarán no sólo a la optimización de la eficiencia de la evaluación sino también a la mejora de su eficacia.

Desde que Sechrest (1963) propusiera inicialmente el concepto de validez incremental, ésta ha sido analizada sobre varias dimensiones: sensibilidad al cambio, eficacia diagnóstica, validez de contenido, validez criterial, diseño de tratamiento y resultado o validez convergente (Haynes y Lench, 2003). Mediante cada una de ellas se ha tratado de contrastar, lo que hoy genéricamente se entiende por *validez incremental*: grado en el que la obtención de una nueva medida mejora la predicción realizada de un criterio a partir de un conjunto de medidas ya dadas (Haynes y Lench, 2003; Husley y Meyer, 2003).

Concretamente, y frente a la validez incremental de las inferencias o de las nuevas medidas generadas, una de las aproximaciones a la

validez incremental se centra en el valor añadido que diferentes fuentes de datos obtenidas a partir de diversos informantes o instrumentos (*incremental validity of testing*) añaden a una ecuación estadística. Este tipo de validez incremental se realiza generalmente sobre la base del análisis de regresión y proporciona evidencia de la validez concurrente o predictiva de las medidas en cuestión. Claramente, como han destacado Husley y Meyer (2003) la contaminación producida entre instrumentos que se administran conjuntamente así como otras formas de sesgos y errores (ej., tipo de análisis estadístico, diseño utilizado, delimitación y fiabilidad del criterio, contaminación artefactual del criterio por los predictores etc...) hacen necesaria una mayor investigación en el ámbito de la validez incremental y la conveniencia de replicar los resultados ya existentes. A esto es necesario sumar la escasa literatura existente sobre el grado en el que varios métodos y fuentes de evaluación contribuyen a la validez incremental para propósitos específicos, dado que, en la mayoría de las situaciones, el uso en investigación de múltiples fuentes y métodos responde más a cuestiones teóricas que a cuestiones relacionadas con la práctica en evaluación.

En población infantil, como ocurre con la adulta, el estudio de la validez incremental en relación con el uso de diferentes métodos y fuentes sigue siendo una asignatura pendiente (Johnston y Murray, 2003). En lo que específicamente se refiere al uso de múltiples fuentes informantes, algunos trabajos apoyan el valor incremental de informantes adultos frente a la información aportada por los propios niños en relación con problemas exteriorizados, como la agresión (Loeber, Green, Lahey y Stouthamer-Loeber, 1991). Sin embargo, el uso de múltiples fuentes procedentes de adultos no siempre añade validez incremental al uso de una sola fuente (Biederman, Keenan y Faraone, 1990), aunque claramente esto debe ser analizado a la luz de la meta específica de la evaluación (ej., discriminar tipos de desórdenes, realizar un diagnóstico, etc...) y de las características particulares de la muestra de estudio (ej., edad, sexo, tipo de alteración presente, etc...) (Johnston y Murray, 2003). Por otra parte, y paradójicamente, abundante investigación aporta más evidencia a la

discrepancia que a la similitud entre las diversas fuentes informantes cuando se trata de evaluar algún contenido relativo a la población infantil (Achenbach y McConaughy, 1987; Cole, Martin, Powers y Truglio, 1996; Del Barrio, 1990, 2002; Epkins, 1993; Jané, Araneda, Valero y Doménech, 2000; Klein, 1991; Silva, 1995).

En el presente trabajo nos disponemos a estudiar la validez incremental que posee la agresión infantil informada por diversas fuentes sobre la predicción de la hostilidad y comunicación maternas. Es clara la asociación entre la hostilidad materna (Bierman y Smoot, 1991; Chen, Huanfeng, Wang y Cen, 2001; Ge, Best, Conger y Simons, 1996; Muris et al., 2003; Muris et al., 2004; Roelofs et al., 2006; Stormshak et al., 2000) y la ausencia de comunicación y afecto positivo con las conductas exteriorizadas de los hijos, en general, y la conducta agresiva, en particular (Chen et al, 2001; Conger et al., 1994; Stice y Barrera, 1995; Rubin y Burgess, 2002; Scaramella, Conger y Simona, 1999). Aunque esta relación ha sido evidenciada en ambas direcciones (Collins et al., 2001; Holden y Miller, 1999; Kerr, Stattin, Biesecker y Ferrer-Wreder, 2003), una de las hipótesis unilaterales bien fundamentadas sobre este tópico es cómo el grado de conflictividad y desajuste infantil propicia y favorece la aparición del rechazo materno, de su hostilidad, y falta de comunicación y afecto (Morris et al., 2002; Reitz et al., 2006; Stice y Barrera, 1995).

En estos estudios, tanto profesores, compañeros, como el propio sujeto se han mostrado buenos informadores de la conducta agresiva infantil (Del Barrio, López y Moreno, 2001; Caprara et al., 1993; Pastorelli, et al., 1997). Por otra parte, los propios hijos son fuentes de información fiables de las pautas de crianza parentales (Carrasco, Holgado y del Barrio, 2007; Morris, Silk, Steinberg, Sessa, Avenevoli y Essex, 2002; Roa y Del Barrio, 2002; Samper, Cortés, Mestre, Náchter y Tur, 2005); sin embargo, poco puede decirse sobre qué fuente predice mejor la conducta parental (compañeros, profesores o el propio niño) cuando la agresión del niño es utilizada como predictor de la misma. En un intento de aportar evidencias empíricas a esta cuestión, se analiza la validez incremental de la agresión infantil evaluada por diversas fuentes sobre la hostilidad y la comunicación maternas.

## Método

### Muestra

Inicialmente, participaron 1116 sujetos con previo consentimiento autoinformado por parte de sus padres y del equipo directivo de la institución escolar correspondiente. Considerando la muestra longitudinalmente, se obtuvieron registros de un total de 525 sujetos (41,3 % varones) en tres ocasiones de medida. La muestra fue extraída de un conjunto de 10 colegios seleccionados aleatoriamente de la Comunidad de Madrid. La media de edad fue de 12.72 (*D.T.* = 1.57) con valores que oscilaron de 9 a 15 años a lo largo de los tres momentos de evaluación: T1 (*M* = 11.10; *DT* = 1.55), T2 (*M* = 12.17; *DT* = 1.61), T3 (*M* = 13.19; *DT* = 1.71). El nivel económico de las familias de pertenencia, según indicadores del Índice de Hollingshead (1975), fue de niveles altos en un 10.7%, medio-bajo en un 41.4% y bajo en un 2.4%. La mayoría de las familias (87.8%) estaba constituida por padres y madres y la media del número de hermanos fue de dos.

### Instrumentos

Para la obtención de las diferentes medidas se utilizaron cuatro instrumentos:

– *Inventario del Comportamiento Parental para Niños* (CRPBI; Schaefer, 1965; adaptación española Samper et al., 2006; Carrasco, Del Barrio y Holgado, 2007).

Este instrumento está diseñado para evaluar el comportamiento parental que perciben los hijos, tanto en su relación con el padre como en la relación con su madre. Consta de 52 ítems referidos a actividades cotidianas relacionadas con la crianza. Los mismos elementos se cumplimentan para el padre y para la madre siguiendo una escala de mayor o menor presencia de esa conducta o actitud (3, «siempre»; 2 «algunas veces»; y 1, «nunca»). Para la realización del presente trabajo se utilizaron las dimensiones obtenidas por Carrasco et al. (2007) y referidas sólo a la figura materna. Concretamente, las dimensiones de segundo orden utilizadas como criterio fueron la de Hostili-

dad/Negligencia materna (Alfa de Cronbach = 0.86) y Comunicación/Afecto materna (Alfa de Cronbach = 0.70). El resto de escalas del instrumento tienen índices de fiabilidad, tanto para la madre como para el padre, comprendidos entre valores de 0.65 y 0.87.

– *Escala de Agresividad Física y Verbal (AFV)* (Caprara y Pastorelli, 1993; adaptada al castellano por Del Barrio, Moreno y López, 2001).

Esta escala, de origen italiano, consta de 20 ítems que evalúan la conducta de hacer daño a otros física y verbalmente. El formato de respuesta es de tres alternativas (a menudo, algunas veces o nunca) según la frecuencia de aparición de la conducta. Además de una puntuación total de agresividad, posee dos factores: agresividad física (Ej., «pego patadas y puñetazos» o «muerdo a otros») y agresividad verbal (Ej., «digo tacos» o «insulto a mis compañeros») y una escala de control para identificar la tendencia de respuesta. Su aplicación es individual o colectiva y está diseñada para sujetos de entre 7 y 17 años. La duración aproximada para su cumplimentación es de 10-15 minutos

Respecto a las características psicométricas de la prueba, la consistencia interna obtenida mediante el alfa de Cronbach alcanza un alfa de 0.86 (Caprara y Pastorelli, 1993). En población normal española los valores de este índice ha oscilado entre 0.73 (Del Barrio, Mestre, Tur y Samper, 2004) y 0.84 (Del Barrio *et al.*, 2001).

– *Registro de Agresividad Heteroinformada por los compañeros* (versión experimental).

Se trata de un registro de doble entrada con filas y columnas. En cada fila aparece el número de lista de clase correspondiente a cada alumno (tantas como alumnos haya) y en las columnas cada uno de los niveles de la escala Likert comprendida entre 1 y 7. El registro consta de un ítem, *¿en qué medida es agresivo este alumno?*, que se aplica a cada compañero. Los sujetos han de asignar a cada compañero una puntuación comprendida entre 1 «nada agresivo» y 7 «muy agresivo». Previamente a la asignación de la puntuación se describía un prototipo de sujeto agresivo, para que el niño lo tomara como referencia para decidir en qué grado cada compañero era o no agresivo. El prototipo se descri-

bía como, «es un/a niño/a que rompe cosas y pega patadas cuando se enfada; insulta a los compañeros si no hacen lo que dice; «se enfrenta con los profesores en clase» o «se pelea con todos».

Para facilitar el recuerdo del número de lista, se escribía en la pizarra los números y nombres que correspondían a cada alumno en el momento de la evaluación. Finalmente, se computaba por cada alumno la media de todas las puntuaciones asignadas por el resto de compañeros. Esto permitió obtener una puntuación sobre la agresividad del sujeto, percibida por sus compañeros.

Análisis previos realizados con los datos obtenidos de la aplicación de esta escala han mostrado que puede ofrecer una buena medida heteroinformada de la agresión. Las correlaciones obtenidas en relación con la agresión autoinformada y la agresión informada por profesores son de 0.34 y 0.51 respectivamente (Rodríguez, 2007).

## **Procedimiento**

Una vez realizada la selección de la muestra y obtenida la autorización para la realización del estudio, se solicitaba la participación voluntaria de los sujetos. La administración de las pruebas se realizó colectivamente por psicólogos entrenados en cada una de las aulas seleccionadas. Los niños fueron evaluados a lo largo de tres años consecutivos (T1, T2, T3) en los que se obtuvieron medidas repetidas de las distintas variables. El presente trabajo es parte de un estudio longitudinal más amplio sobre emociones infantiles. Para los presentes análisis se utilizó como variables predictoras las medidas del Tiempo1 y como criterios la puntuación media de las medidas agregadas procedentes de los tres tiempos de evaluación.

Habitualmente se utilizan dos procedimientos para evaluar la validez incremental en los diseños correlacionales. El primero de ellos, mediante el análisis de las correlaciones parciales y semiparciales Tal y como ha señalado Cohen y Cohen (1983), la validez incremental responde al siguiente esquema: Dadas dos variables predictoras (los tests *A* y *B*) y un criterio *Y*, se cal-

cula la correlación entre cada una de ellas y el criterio, eligiendo aquella cuya correlación con el criterio sea más alta, supongamos el test A ( $r_{Y.A}$ ), a continuación y para ver la validez incremental que supondría utilizar también el test B se calcula la correlación semiparcial eliminando la influencia que pueda estar ejerciendo la relación existente entre los dos tests ( $r_{Y(B.A)}$ ), ya que ambos pueden no ser fuentes independientes de variabilidad sino estar compartiendo una parte de la misma y, por lo tanto, puede existir redundancia entre ambas variables. El aumento experimentado en la correlación con el criterio al introducir la segunda variable predictora será la validez incremental que puede ser determinada obteniendo el cambio en  $R^2$  y su magnitud a partir de diferentes criterios (Cohen, 1988, 1992; Husley y Meyer, 2003; Nunnally y Bernstein, 1994). El segundo de ellos, y quizás el más habitual, ha sido mediante el análisis de regresión por orden jerárquico, el cual permite determinar la contribución de una medida sobre la predicción de un criterio después de que una o más variables hayan entrado en el análisis. La aproximación jerárquica (Thompson, 1995) se ha considerado el método de elección frente al método por pasos, dada la tendencia de este último a capitalizar el error muestral y por tanto, a dificultar la generalización o replicación de los resultados.

De acuerdo con lo anteriormente expuesto se han obtenido, en primer lugar, las correlaciones de orden 0, parciales y semiparciales entre las variables independientes y cada una de las variables las dependientes; y a continuación, se ha ejecutado un análisis de regresión utilizando el método «introducir» en orden jerárquico. Como es sabido, una de las consideraciones más controvertidas a tener en cuenta en el análisis de regresión es el orden de entrada de las variables. Como han sugerido algunos autores (Cohen y Cohen, 1983; Robins, 1987), en el primer bloque del análisis de regresión hemos incluido las variables relevantes de carácter más estable, edad y sexo. En el segundo, tercer y cuarto bloque, se ha añadido, respectivamente, cada uno de los predictores cuya validez incremental se pretendía evaluar: agresividad evaluada por los compañeros, agresividad evaluada por los profesores y agresividad autoinformada. De este modo y dado que las variables inicialmente in-

troducidas (*Edad* y *Sexo*) permanecen constantes a lo largo del proceso de análisis, funcionarán como variables control.

Controlando dichas variables, este procedimiento permite calcular si la incorporación de cada una de las fuentes de información supone un cambio sobre el modelo previo en el poder explicativo de las variables criterio. Como criterios se tomaron dos variables de signo contrario: la hostilidad y la comunicación/afecto maternas evaluadas por el propio niño. Con el fin de compensar el solapamiento artefactual de la fuente niño, compartida en el criterio y en uno de los predictores, ambos criterios fueron estimados a partir de la media de las puntuaciones agregadas a lo largo de los tres momentos de evaluación (T1, T2, T3) realizados, separados cada uno por un año consecutivo.

El conjunto de análisis han sido realizados con el paquete estadístico SPSS versión 15.0 para Windows (SPSS WIN).

## Resultados

La determinación de la significación estadística es dependiente del tamaño muestral y hay grandes dificultades para determinar cómo un incremento estadísticamente significativo tiene alguna utilidad o valor clínico. Hunsley y Meyer (2003) proponen dos opciones para considerar el grado de incremento estadísticamente significativo:

Indirectamente examinando el grado por el que la asociación entre una medida y el criterio es dependiente de la varianza compartida con otras variables en la ecuación de regresión. Lo que Lindenberger y Pötter (1998) denominaron *aproximación sobre los efectos simples compartidos (SOS; Shared over simple effects approach)*. Para evaluar los efectos simples y compartidos en el análisis de regresión: el mayor tamaño relativo del incremento en la validez es el valor más pequeño de los efectos compartidos (SOS).

El SOS no ofrece una métrica absoluta para la evaluación del efecto de la validez incremental pero es una vía apropiada para hallar la significación del incremento respecto de la naturaleza del criterio.

A través del cambio en  $R^2$ . Su magnitud puede ser interpretada a la luz de algún punto de referencia conceptualmente razonable: 1) Cohen (1988, 1992) propone que el tamaño medio de esta relación sea sobre 0.30 para predictores y criterios medidos independientemente, oscilando en un rango de pequeño a medio (desde 0.10 a 0.30); 2) Nunnally y Bernstein (1994) proponen que el incremento en  $R$  generalmente ha de ser más pequeño en el momento en el que un tercer predictor es incluido en una ecuación de regresión; 3) Husley y Meyer (2003) proponen que una correlación semiparcial de 0.15 a 0.20 en el tercer paso indicaría una contribución razonable a la ecuación existente que se traduciría en un ligero incremento de  $R$  aproximadamente de 0.03.

Para ver si el aumento es estadísticamente significativo se puede utilizar también el siguiente estadístico de contraste:

$$F = \left( \frac{N - K - 1}{K - J} \right) \left( \frac{R_{Y.KX}^2 - R_{Y.JX}^2}{1 - R_{Y.KX}^2} \right)$$

Donde:

$N$  = número de sujetos de la muestra

$K$  = número de predictores finales incluidos

$J$  = número de predictores incluidos hasta el paso anterior

$R_{Y.KX}^2$  = correlación múltiple al cuadrado con  $K$  predictores

$R_{Y.JX}^2$  = correlación múltiple al cuadrado con  $J$  predictores

El estadístico de contraste tiene una distribución  $F$  de Snedecor con  $(K-J)$  y  $(N-K-1)$  grados de libertad (Barbero, Vila y Suárez, 2006, pág. 376).

**Tabla 1. Correlaciones de orden 0, parciales y semiparciales entre las variables predictoras y criterio.**

	HOSTILIDAD			COMUNICACIÓN		
	Orden 0	Parcial	Semiparcial	Orden 0	Parcial	Semiparcial
Agresividad informada por:						
Iguales	0.140	0.146	0.145	-0.093	-0.055	-0.054
Profesores	0.107	0.024	0.023	-0.168	-0.157	-0.156
Autoinformada	0.235	0.238	0.233	-0.257	-0.212	-0.208

Realizando una primera aproximación de carácter meramente descriptivo a través del conjunto de correlaciones de orden 0, parciales y semiparciales, encontramos que, la agresión autoinformada es la que más correlaciona tanto con la hostilidad, como con la comunicación materna (ver Tabla 1). En cuanto a hostilidad, si eliminamos la influencia que sobre dicha variable están ejerciendo la agresión informada por los iguales y por el profesor (correlación semiparcial), la magnitud de dicha correlación apenas se resiente. De hecho, de 0.235 pasa a ser 0.233, lo que estaría indicando que ambas variables comparten poca variabilidad. Atendiendo a la correlación parcial, es decir, cuando excluimos el efecto que sobre la correlación entre agre-

sividad auto-informada y hostilidad materna tienen el resto de variables, observamos que los cambios también son mínimos (0.235 a 0.238).

En el caso de las correlaciones entre la agresión informada por distintas fuentes y la comunicación materna, se observa la misma tendencia, es decir, la magnitud más fuerte se da entre la comunicación y la agresión auto-informada.

En términos de los criterios propuestos por Lindenberger y Pötter (1998), el valor más pequeño de los efectos compartidos sería el mayor tamaño relativo del incremento en la validez. En ambos casos, parece ser que la agresión auto-informada es la que comparte menos variabilidad con el resto de predictores, en su

combinación con las variables criterio. Por tanto, en el conjunto de variables predictoras sería la de mayor validez incremental.

Una vez descritas las correlaciones parciales y semiparciales, se analiza si el cambio en  $R^2$  de un modelo a otro es significativo. Es decir, tratamos de determinar si la incorporación al modelo de una nueva variable predictiva (según

distintos informantes) añade información que sea relevante para explicar la variabilidad de la variable dependiente. En este caso, una vez controlada la influencia de las variables sexo y edad, vamos incorporando en el modelo las puntuaciones en agresión de cada una de las fuentes de información para predecir las variables hostilidad y comunicación materna.

**Tabla 2. Regresión jerárquica con el método «introducir» sobre hostilidad, controlando edad y sexo.**

Modelo	$\beta$	$t$	Estadísticos del cambio		
			$R^2$ corregida	Cambio en $R^2$	Cambio en $F_{(g1, g2)}$
Edad	-0.127	-1.612			
Sexo	-0.056	-0.716	0.021	0.003	0.512 <sub>(1,159)</sub>
Edad	-0.151	-1.900			
Sexo	0.004	0.053	0.042	0.021	3.462 <sub>(1,1582)</sub>
Iguales	0.158	1.861			
Edad	-0.146	-1.804			
Sexo	0.003	0.037	0.042	0.000	0.088 <sub>(1,157)</sub>
Iguales	0.145	1.530			
Profesor	0.026	0.297			
Edad	-0.169	<b>-2.137*</b>			
Sexo	0.073	0.844			
Iguales	0.174	1.873	0.096	0.054*	9.340 <sub>(1,1562)</sub>
Profesor	-0.036	-0.405			
Auto-informada	0.248	<b>3.056**</b>			

$\beta$  = Coeficientes estandarizados

\*  $p < 0,05$ ; \*\*  $p < 0,01$

Los resultados de la tabla 2, indican que la incorporación de las puntuaciones de la agresividad informada por los iguales no resulta significativa. Este mismo resultado lo encontramos para la información que puedan aportar los profesores. Sin embargo, cuando hemos incorporado en el modelo la agresión autoinformada, sí que hemos encontrado un cambio significativo en el poder explicativo del modelo.

Es decir, la agresión autoinformada tiene un valor predictivo incremental sobre la agresión heteroinformada, ya sea por los compañeros o por los profesores, respecto de la hostilidad materna. Por tanto, el uso de la fuente heteroinformada no sólo no es informativa de la conducta hostil materna, sino que no tiene valor predictivo adicional respecto de la fuente autoinformada. En cambio, cuando la agresividad es infor-

mada por el propio niño, la conducta agresiva predice significativamente la conducta hostil materna. Aunque efectivamente, el cambio en  $R^2$  está lejos de los criterios establecidos por Cohen (1988,1992).

Es también destacable, que la variable edad adquiera un valor significativo en este último modelo en el que se incorpora la agresión autoinformada. Este dato, indica la posible interacción entre edad y agresión del niño como predictores de la hostilidad materna.

Respecto a la variable comunicación (ver tabla 3), al igual que en hostilidad, la información aportada por los iguales no incrementa significativamente el poder explicativo del modelo; sin embargo, y a diferencia del anterior, la inclusión de la información aportada por los pro-

fesores sí resulta significativa. Finalmente, es la información aportada por los propios niños la que incrementa de manera significativa el poder explicativo del modelo. Es decir, el profesor frente a los compañeros, tiene valor predictivo, pero cuando añadimos la fuente autoinformada al modelo la fuente profesores queda desplazada. Lo que indica que hay una información redundante entre la aportada por el profesor y el niño, pero en la que prevalece la del propio niño. Por tanto, frente a los iguales el profesor predice mejor la conducta comunicativa materna, sin embargo cuando son utilizadas las tres fuentes, es la del propio niño la que tiene mayor valor predictivo. Claramente es la fuente autoinformada la que presenta un claro «valor añadido» frente a la del profesor y la de los compañeros.

**Tabla 3. Regresión jerárquica con el método «introducir» sobre comunicación, controlando edad y sexo.**

Modelo	$\beta$	$t$	Estadísticos del cambio		
			$R^2$ corregida	Cambio en $R^2$	Cambio en $F_{(gl1, gl2)}$
Edad	-0.065	-0.819			
Sexo	0.097	1.229	0.003	0.003	1.509 <sub>(1,159)</sub>
Edad	-0.056	-0.698			
Sexo	0.074	0.866	0.003	0.000	0.473 <sub>(1,158)</sub>
Iguales	-0.059	-0.687			
Edad	-0.087	-1.069			
Sexo	0.084	0.981			
Iguales	0.025	0.265	0.015	0.012*	3.985 <sub>(1,157)</sub>
Profesor	-0.177	<b>-1.996*</b>			
Edad	-0.066	-0.825			
Sexo	0.022	0.248			
Iguales	-0.001	-0.007	0.054	0.039*	7.357 <sub>(1,156)</sub>
Profesor	-0.121	-1.358			
Auto-informada	-0.222	<b>-2.712*</b>			

$\beta$  = Coeficientes estandarizados

\*  $p < 0,05$ ; \*\*  $p < 0,01$

## Discusión y conclusiones

Los principales resultados de este trabajo muestran que la información aportada por los propios niños sobre su conducta agresiva es la que mejor predice la hostilidad y la comunicación maternas. La fuente autoinformada, procedente del propio niño, frente a la heteroinformada, procedente de compañeros y profesores, fue la que incorporó cierto valor añadido a la predicción del comportamiento materno. Este dato, además de confirmar la relación bien establecida entre el comportamiento agresivo del niño y la hostilidad (Bierman y Smoot, 1991; Chen, Huanfeng, Wang y Cen, 2001; Ge et al., 2006; Muris et al., 2003; Muris et al., 2004; Roelofs et al., 2006; Stormshak et al., 2000) y disminución de la comunicación maternas (Chen et al., 2001; Conger et al., 1994; Stice y Barrera, 1995; Rubin y Burgess, 2002; Scaramella, Conger y Simona, 1999), apoya los trabajos previos en los que el propio niño se considera una fuente informante fiable y válida de su propia conducta, particularmente de la agresiva (Del Barrio, López y Moreno, 2001; Caprara et al., 1993; Pastorelli, et al., 1997). Pero quizás, lo que resultó más novedoso de este trabajo es el apoyo que los presentes datos conceden a la superioridad predictiva de la fuente informante niño sobre la fuente profesor o compañero. La escasez de trabajos en esta línea no permite obtener comparaciones con trabajos previos, lo que anima a la replicación de los presentes resultados con otro tipo de muestras y medidas.

Como bien advierten algunos autores (Johnston y Murray, 2003), la estimación de la validez incremental de una medida sobre otra, ha de ser analizada desde la especificidad y características de la evaluación que se realiza. Desde esta perspectiva, podemos afirmar que la información del niño sobre su conducta agresiva predice en mayor medida la conducta hostil y comunicativa de la madre. La combinación con otras fuentes no mejora la predicción de dicha conducta y la optimización de la eficiencia evaluadora, cuando se trate de estos criterios, quizás pase por excluir a los compañeros y profesores como informantes. Atendiendo a la especificidad a la que aludimos, esta validez incremental de la fuente niño sobre la de profesores y compañeros, al menos ha de ser defendida

cuando la edad de los sujetos informantes sea de 9 a 15 años, se informe sobre la conducta agresiva del propio niño y se predigan los hábitos de crianza hostiles y comunicativos de la madre. No obstante, es importante diferenciar la validez incremental derivada del uso de instrumentos combinados, como es el presente caso, de la validez incremental de las inferencias. Esta última, relacionada, en mayor medida, con la toma de decisiones y la predicción de las variables criterio. En este sentido, es relevante tener presente que la precisión de las predicciones aumenta cuando las fuentes de datos no proveen información redundante (Tsujiyamoto, Hamilton y Berger, 1990). La evidencia de la validez incremental de una medida (fuente o instrumento) (*validity incremental of testing*), no garantiza necesariamente una mejora en la toma de decisiones para las que la medida es usada (*Incremental validity of test informed inference*) (Hunsley y Meyer, 2003). Por tanto, en el caso que nos ocupa, el uso de otras fuentes distintas a la del niño, aunque redundantes, puede estar justificado para la mejora de las inferencias y la toma de decisiones, pero no para la mejora en la precisión y predicción del criterio objeto de evaluación. Esta circunstancia, entronca directamente con la noción de validez consecucional (Messick, 1994). En este sentido, entendemos que si bien no garantiza una mejora en la toma de decisiones, el hecho de tomarlas basándonos en fuentes o instrumentos con mayor validez incremental potencia el buen uso de los tests y de la evaluación de cara a las consecuencias que de ello puedan derivarse. Así pues, estos datos no restan importancia a la defensa de la evaluación multimétodo y multifuente, tan sólo realzan la necesidad de optimizar el uso combinado y empíricamente sustentado de dicho proceder según los objetivos perseguidos en la evaluación.

Una de las limitaciones de este trabajo es lo que dentro de la contaminación artificial del criterio se ha denominado solapamiento artificial de la fuente (*source overlap artifact*) existente cuando el predictor y el criterio son evaluados por una misma fuente. En este caso, la asociación empírica entre el predictor y el criterio es función de una tercera variable no medida (fuente), que influencia a ambas medidas y revierte en una artificial y elevada estimación del

predicador sobre el criterio. Una posible solución para ello es, precisamente, el empleo de matrices multirrasgo-multimétodo, que en este caso sería multifuente. En modelos de ecuaciones estructurales, análogamente a la estimación del efecto del método (*method factor*), se trataría de determinar el efecto de la fuente. En nuestro estudio, lo hemos tratado de soslayar mediante la agregación de información para la obtención del criterio (información en varias ocasiones longitudinalmente, con varios ítems o estímulos, varios instrumentos o métodos, varias fuentes de información), intentando de esta forma aumentar la fiabilidad y validez de la información agregada (Epstein, 1980, 1983).

Desde que Campbell y Fiske (1959) propusieran las matrices multirrasgo-multimétodo, la investigación se ha centrado en la medida de distintos constructos con distintos métodos, pero no con distintos informantes, siendo esta última una cuestión crucial para la evaluación psicológica, en general, y para el concepto de validez en particular. Por ello, entendemos que el uso de distintos informantes debería ser analizado desde un marco teórico más amplio, y precisamente el uso de matrices multirrasgo-multimétodo (*multifuente*) puede ser apropiado

para ello y aportar resultados muy interesantes. De esta forma, y a título de ejemplo, el concepto tradicional de validez convergente podría no limitarse al uso de distintos métodos, sino también al uso de distintas fuentes.

Otra limitación importante de nuestro trabajo es el escaso valor predictivo de las variables independientes utilizadas. Los resultados distan de los criterios establecidos en la relevancia del cambio en  $R^2$ . Por ello, debemos seguir indagando sobre las variables medidas para, en la medida de lo posible, enriquecer los modelos explicativos.

Futuros trabajos necesitarían abordar la validez incremental proporcionada por el uso de diversos métodos, el coste diferencial de diferentes informantes, el valor de diferentes informantes según la etapa del proceso de evaluación en la que nos encontremos, las características de los informantes que llevan a la producción de información única (ej., el número de contactos que mantienen las fuentes informantes con el niño), la optimización de la combinación de información discrepante procedente de diversas fuentes, y el estudio de la influencia que tiene las inconsistencias recogidas en el objetivo de la evaluación.

## Referencias Bibliográficas

- Achenbach, T.M. y McConaughy, S.H., y Howell, C.T. (1987). Chile/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, 101, 213-232.
- Barbero, M<sup>a</sup>.I., Vila, E. y Suárez, J.C. (2006). *Psicometría*. Madrid: UNED.
- Biederman, J., Keenan, K., y Faraone, S. V. (1990). Parent-based diagnosis of attention deficit disorder predicts a diagnosis based on teacher report. *Journal of the American Academy of Child and Adolescent Psychiatry*, 29, 698-701.
- Bierman, K.L. y Smoot, D.L. (1991). Linking family characteristics with poor peer relations: the mediating role of conduct problems. *Journal of Abnormal Child Psychology*, 19, 341-356.
- Campbell, D.T. y Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56 (2), 81-105.
- Caprara, G.V. y Pastorelli, C. (1993). Early emotional instability, prosocial behaviour and aggression: some methodological aspects. *European Journal of Personality*, 7, 19-36.
- Carrasco, M. A., Holgado, P. y Del Barrio M. V. (2007). Análisis de la estructura del cuestionario de comportamiento parental para niños (CRPBI) en población española. *Revista Iberoamericana de Diagnóstico y Evaluación Psicológica*, 24, 95-120.
- Chen, X., Wu, H., Wang, I. y Cen, G. (2001). Parenting practices and aggressive behavior in Chinese children. *Parenting: Science and Practice*, 1, 159-184.
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2<sup>nd</sup>. Ed.). Hillsdale, NJ: Erlbaum.

- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.
- Cohen, J. y Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioural sciences*. New York: Erlbaum.
- Cole, D.A., Martin, J.M., Powers, B. y Truglio, R. (1996). Modeling causal relations between academic and social competence and depression: a multitrait-multimethod longitudinal study of children. *Journal of Abnormal Psychology*, 105, 258-270.
- Collins, W. y Russell, G. (1991). Mother-child and father-child relationships in middle childhood and adolescence: a developmental analysis. *Developmental Review*, 11, 99-136.
- Conger, R.D., Ge, X., Elder, G., Lorenz, F. y Simons, R.L. (1994). Economics stress, coercitive family process, and developmental problems of adolescents. *Child Development*, 65, 541-561.
- De los Reyes, A. y Kazdin, A.E. (2004). Measuring informant discrepancies in clinical child research. *Psychological Assessment*, 16, 330-334.
- Del Barrio, M.V. (2002). Método de la evaluación psicológica y su aplicación al niño y el adolescente. En M.V. del Barrio (Coord.), *Evaluación psicológica en la infancia y la adolescencia* (pp. 55-88). Madrid: UNED.
- Del Barrio, V. (1990). Situación actual de la evaluación de la depresión infantil. *Evaluación psicológica*, 6, 171-209.
- Del Barrio, V., Moreno, C. y López, R. (2001). Evaluación de la agresión, y la inestabilidad emocional en niños españoles: su relación con la depresión. *Clínica y Salud*, 12, 34-50.
- Epkins, C. (1993). A preliminary comparison of teacher ratings and child self-report of depression, anxiety, and aggression in inpatient and elementary school samples. *Journal of Abnormal Child Psychology*, 21, 649-661.
- Epstein, S. (1980). The stability of behaviour: II. Implications for psychological research. *American Psychologist*, 35, 790-806.
- Epstein, S. (1983). Aggregation and beyond: some basic issues on the prediction of behaviour. *Journal of Personality*, 51, 360-392.
- Fernández-Ballesteros, R. (1994b). Psychological Assessment. *Internacional Journal of Applied Psychology*, 43, 157-174.
- Garb, H.N. (2003). Incremental Validity and the assessment of psychopathology in adults. *Psychological Assessment*, 15, 508-520.
- Ge, X., Best, K., Conger, R. y Simons, R. (1996). Parenting behaviours and de occurrence and co-occurrence of adolescent depressive symptoms and conduct problems. *Developmental Psychology*, 32, 717-731.
- Haynes, S. y Lench, H. (2003). Incremental validity of new clinical assessment measures. *Psychological Assessment*, 15, 456-466.
- Holden, G. y Miller, P. (1999). Enduring and different: A meta-analysis of the similarity in parents' child rearing. *Psychological Bulletin*, 125, 223-254.
- Hunsley, J. (2003). Introduction to the special section on incremental validity and utility in clinical assessment. *Psychological Assessment*, 15, 443-445.
- Hunsley, J. y Mash, E.J. (2005). Introduction to the special section on developing guidelines for the evidence based assessment (EBA) of adult disorders. *Psychological Assessment*, 17, 251-255.
- Husley, J. y Meyer, G. (2003). The incremental validity of psychological testing and assessment: conceptual, methodological, and statistical issues. *Psychological Assessment*, 15, 466-455.
- Jané, M., Araneda, N., Valero, S. y Doménech, E. (2000). Evaluación de la sintomatología depresiva del preescolar: correspondencia entre los informes de padres y de maestros. *Psicothema*, 12, 212-215.
- Johnston, Ch. y Murray, C. (2003). Incremental validity in the psychological assessment of children and adolescents. *Psychological Assessment*, 15, 496-507.
- Kerr, M., Stattin, H, Biesecker, G y Ferrer-Wreder, L. (2003) Relationships with parents and peers in adolescence. Lerner (Ed). *Handbook of psychology: Developmental psychology*, Vol. 6 (pp. 395-419). New York: John Wiley and Sons.
- Klein, R.G. (1991). Parent-child agreement in clinical assessment of anxiety and other psychopathology: a review. *Journal of Anxiety Disorders*, 5, 187-198.
- Loeber, R., Green, S.M., Lahey, B.B., y Stouthamer-Loeber, M. (1991). Differences and similarities between children, mothers and teachers as infor-

- mants on disruptive behaviour disorders. *Journal of Abnormal Child Psychology*, 19, 75-95.
- McFall, R. (2005). Theory and utility-Key themes in evidence based assessment: comment on the special section. *Psychological Assessment*, 17, 312-323.
- Messick, S. (1994). Foundations of validity: Meaning and consequences in psychological assessment. *European Journal of Psychological Assessment*, 10(1), 1-9.
- Meyer, G.J. y Archer, R.P. (2001). The hard science of Rorschach research: what do we know and do we go?. *Psychological Assessment*, 13, 486-502.
- Morris, A., Silk, J., Steinberg, L., Sessa, F., Avenevoli, S. y Essex, M. (2002). Temperamental vulnerability and negative parenting as interacting of child adjustment. *Journal of Marriage and Family*, 64, 461-471.
- Muris, P., Meesters, C. y van den Berg, S. (2003) Internalizing and externalizing problems as correlates of self-reported attachment styles and perceived parental rearing in normal adolescent. *Journal of Child and Family Studies*, 12, 171-183.
- Muris, P., Meesters, C., Morren, M. y Moorman, L. (2004). Anger and hostility in adolescents: relationships with self-reported attachment style and perceived parental rearing styles. *Journal of Psychosomatic*, 57, 257-264.
- Nunnally, J.C. y Bernstein, I.H. (1994). *Psychometric theory* (3rd Ed.). New York: Wiley.
- Pastorelli, C., Barbaranelli, C., Cermak, I., Rozsa, S. y Caprara, G.V. (1997). Measuring emotional instability, prosocial behavior and aggression in preadolescents: a cross-national study. *Personality and Individual Differences*, 23, 4, 691-703.
- Reitz, E., Dekovic, M. y Meijer, A. (2006) Relations between parenting and externalizing and internalizing problem behaviour in early adolescence: Child behaviour as moderator and predictor. *Journal of Adolescence* 29, 419-436.
- Roa, L. y Del Barrio, V. (2002). Cuestionario de percepción de crianza para niños y adolescentes. *Psicología Educativa*, 8, 1,37-51.
- Robins, C.J. (1987). On interpreting results of multiple regression procedures: A cautionary note for researchers and reviewers. *Cognitive Therapy and Research*, 11, 705-708.
- Roelofs, J., Meesters, C., Huurne, M., Bamelis, L. y Muris P. (2006). *Journal of Child and Family Studies*, 15, 331-344.
- Rubin, K. y Burgess, K. (2002). Parents of aggressive and withdrawn children. En Bornstein, Marc H (Ed). (2002). *Handbook of parenting: Children and parenting*, vol. 1 (pp. 383-418). New York: Lawrence Erlbaum Associates.
- Samper, P Cortés, M., Mestre, V., Náchter, M. y Tur, A. (2006). Adaptación del Child's Report of Parent Behavior Inventory (Schaefer, 1965) a población española. *Psicothema*, 18, 263-271.
- Scaramella, L., Conger, R. y Simons, R. (1999). Parental protective influences and gender-specific and externalizing problems differences. *Journal of research on adolescence*, 9, 111-141.
- Sechrest, L. (1963). Incremental validity: a recommendation. *Educational and Psychological Measurement*, 23, 153-158.
- Silva, F. (1995). Cuestiones introductorias. En F. Silva (Ed.), *Evaluación psicológica en niños y adolescentes* (24-42). Madrid: Síntesis.
- Stice, E. y Barrera, M. (1995). A longitudinal examination of the reciprocal relations between perceived parenting and adolescents' substance use and externalizing behaviours. *Developmental Psychology*, 31, 2, 322-334.
- Stormshak, E., Bierman, K., McMahon, R. y Lengua, L. (2000). Parenting practices and child disruptive behaviour problems in early elementary school. *Journal of Clinical Child Psychology*, 29, 1.
- Thompson, B. (1995). Stepwise regression and stepwise discriminant analysis need not apply here. A guidelines editorial. *Educational and Psychological Measurement*, 55, 525-534.