

## Valoración de «Filosofía para Niños»: un programa de enseñar a pensar

Félix García-Moriyón, Roberto Colom\*, Santos Lora\*\*, María Rivas\*\*\* y Vicente Traver\*\*\*\*

IES Avenida de los Toreros, \* Universidad Autónoma de Madrid, \*\* IB de Boadilla del Monte,

\*\*\* IB Julio Palacios de San Sebastián de los Reyes y \*\*\*\* IB Móstoles

El programa de enseñar a pensar aplicado internacionalmente «Filosofía para Niños» (FpN) ha sido desarrollado por Mathew Lipman. Su diseño intenta promover en el alumno tanto las habilidades básicas de pensamiento crítico y creativo como una serie de componentes afectivos. El enfoque de Lipman es filosófico y su base teórica está en Peirce, Mead y Dewey. El presente estudio valora el impacto del programa en tres institutos de la Comunidad de Madrid, analizando los patrones de cambio en una serie de variables de un grupo experimental que recibe el programa y de un grupo de control que no lo recibe, durante un curso académico. Los resultados sugieren una mejora estadísticamente significativa del grupo experimental en una medida de la capacidad cognitiva general, pero no en las medidas de las capacidades cognitivas específicas ni en una serie de componentes básicos de la personalidad.

*Evaluation of «Philosophy for Children»: a program of learning thinking skills.* Philosophy for Children (P4C) is a program developed originally by Mathew Lipman. P4C tries to improve several thinking skills as well as the students' critical and creative thinking. Their conative traits are also addressed. Lipman's perspective is philosophical and his theoretical basis can be found in Peirce, Mead and Dewey. The present study evaluates the impact of P4C in three colleges at the Comunidad de Madrid. The pattern of changes in several cognitive and conative variables were analysed following the application of the program to experimental and control groups over a whole school year. The results suggests an improvement of the experimental group in general cognitive ability, but neither in specific cognitive abilities nor in personality.

Existen varios estudios que muestran el efecto beneficioso que ir a la escuela produce en el desarrollo cognitivo de los niños (Brody, 1992, 1997; Ceci y Williams, 1997; Colom, 1998). Sin embargo, las variables escolares concretas que influyen sobre el rendimiento cognitivo son por ahora desconocidas (Neisser et al., 1996). El mejor predictor del rendimiento escolar son las puntuaciones en los tests estandarizados de inteligencia. La correlación promedio es de  $r=.5$ . No existe ningún otro predictor con una correlación tan alta. Además, el rendimiento en los tests presenta un elevado grado de estabilidad: la correlación entre el rendimiento a los 6 y a los 12 años de edad es de  $r=.89$ ; entre los 6 y los 18 años es de  $r=.77$ . Según Detterman y Thompson (1997) esta importante relación señala la necesidad de comprender por qué se produce y sugiere que el diseño de programas de investigación dirigidos a comprender las causas de las diferencias de rendimiento en los tests estandarizados de inteligencia, puede aportar claves relevantes para la mejora del rendimiento educativo en general. Es posible entrenar a los alumnos para mejorar su nivel en, por ejemplo, matemáticas. Sin embargo, esa mejora no tenderá a generalizarse a otras materias del curriculum (Hunt, 1995). Por consiguiente,

puede ser relevante tratar de acrecentar las aptitudes cognitivas que pueden influir en el rendimiento general de los alumnos en la escuela (Nickerson et al., 1994; Snow, 1989). En este sentido, una parte significativa de los programas de mejora actualmente disponibles ponen como criterio de calidad la comprobación del incremento de rendimiento en tests estandarizados de inteligencia (Nickerson et al., 1994; Prieto y Pérez, 1993).

Los alumnos desarrollan sus capacidades a través de la educación. Pero, según Detterman y Thompson (1997) no todos los alumnos parecen sacar el mismo provecho. La educación puede estar dirigida a aumentar la media de rendimiento o reducir las diferencias de rendimiento (varianza). Esta distinción es relevante, puesto que si la educación considera un método estandarizado de enseñar, el resultado puede ser un incremento promedio de rendimiento y un aumento de las diferencias de rendimiento (Colom, 1997).

Prieto y Pérez (1993) consideran que un programa de enseñar a pensar o de mejora de la inteligencia «ha de ser sensible a las diferencias individuales» (p. 116). Asimismo, en el programa de enriquecimiento instrumental de Feuerstein se contemplan las diferencias individuales y el ritmo de aprendizaje de cada individuo (Prieto y Pérez, 1993, p. 32). Pero ¿cuáles son las características de los alumnos a las que adaptar la educación? Existen distintas perspectivas. Así por ejemplo, está la posición clásica de la psicometría, las teorías cognitivas de autores como Sternberg o Gardner, y las teorías del desarrollo basadas en el enfoque de Piaget.

Neisser et al. (1996) destacan que la perspectiva más empleada en la práctica y la que ha generado un mayor cúmulo de evidencia empírica, es la psicométrica.

El presente estudio analiza si se producen cambios en el estudiante a raíz del empleo de un programa de mejora de las habilidades básicas de pensamiento denominado Filosofía para Niños (FpN), desarrollado por Mathew Lipman y aplicado actualmente a nivel internacional (Lipman et al., 1992). No se estudian, por tanto, las ganancias a largo plazo.

Este programa se introdujo en España en 1985. Está diseñado para contribuir al desarrollo de las habilidades básicas de un pensamiento crítico y creativo. En el modelo de pensamiento adoptado por el programa se consideran otras habilidades que suelen incluirse en la dimensión afectiva. El programa pretende promover el desarrollo de una serie de habilidades básicas de pensamiento y se asume que la aplicación del programa contribuye a desarrollar las habilidades básicas y a mejorar el aprendizaje y el rendimiento académico en diversas áreas curriculares. Esas habilidades básicas están orientadas a que el alumno desarrolle un pensamiento crítico y creativo (un pensamiento complejo de alto nivel).

Según Lipman (Lipman et al., 1992; 1997) los niños son filósofos por naturaleza. El programa pretende convertirse en una ayuda para que los niños piensen por sí mismos. La finalidad es el desarrollo y refuerzo de las siguientes habilidades básicas: (a) sacar inferencias perceptuales, lógicas y causales; (b) hacer asociaciones y analogías; (c) formular hipótesis; (d) efectuar distinciones significativas; (e) considerar las alternativas; (f) buscar razones, explicaciones y supuestos.

Si bien el programa guarda un cierto parentesco con otros programas semejantes encaminados a desarrollar las habilidades cognitivas fundamentales, podemos señalar cinco características que le diferencian de los demás: (a) es un programa de metacognición: plantea que la mejor manera de mejorar las habilidades de pensamiento es hacer que los niños piensen en su propio pensamiento; (b) es un programa de filosofía: las habilidades propias de una discusión filosófica son adecuadas para ayudar a los niños a desarrollar ese tipo de habilidades; (c) es un programa a largo plazo: considera que una intervención pedagógica significativa se debe extender a todos los años de la escolarización obligatoria, y es por eso por lo que ofrece materiales curriculares para poder trabajar con niños desde los 4 hasta los 18 años; (d) es un programa integral: no sólo busca el desarrollo cognitivo, sino también el afectivo, y presta atención a dimensiones cognitivas, éticas y estéticas; y (e) muestra una preocupación social y política: más que estar interesado en la mejora del rendimiento académico, el programa considera que el desarrollo de este tipo de habilidades es condición necesaria para poder alcanzar sociedades democráticas.

Los materiales que emplea el programa son novelas que los alumnos deben leer en el aula siguiendo unas reglas básicas de investigación en comunidad. Estas novelas van acompañadas de manuales de enseñanza para profesores, diseñados para facilitar una aplicación adecuada del programa. Además de ello, el profesor recibe habitualmente un seminario de formación con profesores ya expertos en su aplicación.

Nickerson et al. (1994) sintetizan los escasos estudios que se han llevado a cabo sobre la valoración del impacto del programa FpN en el desarrollo cognitivo de los alumnos. Posiblemente el estudio más importante sobre el impacto del programa haya sido el realizado por el *Educational Testing Service* (ETS) y en el que se

trató de averiguar si los alumnos mejoraban en los campos de razonamiento abordados por el programa (extraer inferencias formales e identificar errores, descubrir alternativas y posibilidades, proporcionar razones y explicaciones), en su fluidez de ideas, en su preparación académica, en lectura y en matemáticas. Los resultados apoyaron la conclusión de que el programa Filosofía para Niños era eficaz para mejorar el rendimiento intelectual de los alumnos. La siguiente es la conclusión de Nickerson et al. (1994): «nuestra valoración del programa es positiva (...) sospechamos que al tomar en serio a los alumnos como pensadores cuyas ideas merecen ser compartidas y comprendidas, el programa tiene un efecto positivo sobre las actitudes de los alumnos, un efecto que puede resultar difícil de medir (...) sospechamos que los resultados obtenidos en clase pueden depender en gran medida de la habilidad del profesor, pero éste es un comentario que podría hacerse prácticamente sobre todos los programas que hemos considerado en este libro» (p. 329).

En nuestro país se realizó una investigación en 1989 con un grupo de 200 alumnos (García Moriyón et al., 1989); se comprobó entonces que el alumnado mejoraba en inteligencia general (test de Raven) y en lectura comprensiva. Posteriormente se han realizado otras dos investigaciones sobre aspectos concretos (Miranda, 1990; Ferrer, 1991). Con todo, no suponen evidencia suficiente sobre el impacto del programa. Por ello, en el presente estudio intentamos comprobar si, en efecto, tal y como apunta el estudio del ETS, el programa produce un efecto beneficioso sobre los alumnos.

## Método

### *Participantes*

La muestra original del estudio estuvo compuesta por 250 alumnos de la ESO en 3 Institutos de la Comunidad de Madrid. La media de edad fue de 14.7 años (desviación típica = 0.86). A los 250 alumnos de la muestra original se les aplicó una batería de tests en octubre de 1995. Sin embargo, la muestra original se redujo significativamente en la aplicación de la batería en mayo de 1996 (retest), pasando, dependiendo del test, a una muestra de entre 115 y 175 alumnos. Una parte importante de la reducción se debió a los problemas de coordinación horaria y de logística planteados en uno de los centros. Los otros motivos de la reducción fueron: (a) la no disponibilidad de datos para todos los alumnos en todos los tests, puesto que los distintos tests no se aplicaban en la misma sesión; (b) la participación en la investigación era plenamente voluntaria, de modo que algunos de los alumnos participantes no acudieron a alguna de las sesiones; (c) la necesidad de descartar los datos de algunos alumnos por mostrar puntuaciones sospechosas; (d) el alumno realizó marcas inapropiadas en las hojas de respuestas que impidieron su procesamiento óptico. Ello ha producido una desigual distribución de alumnos según el test de la batería.

A pesar de estos problemas, bastante frecuentes en los estudios de carácter longitudinal (Schaie, 1996), la muestra de la fase de test y del retest incluye un número de alumnos suficiente para que las comparaciones sean informativas. Además, los análisis estadísticos que se han llevado a cabo muestran que ésta reducción muestral no es relevante (p.e. las diferencias en la fase de test entre el grupo control y el grupo experimental no son estadísticamente significativas en ninguna de las medidas criterio).

### Medidas y procedimiento

Los estudios sobre la valoración del impacto de un programa de mejora de la inteligencia, suelen realizar medidas pretest-postest para observar posibles diferencias antes y después de la intervención (Calero 1997; Nickerson et al., 1994; Prieto y Pérez, 1993). El tipo de medidas que se emplean para valorar programas tan conocidos como el de enriquecimiento instrumental o el Harvard, suelen ser test de inteligencia y razonamiento como el Dominós, la batería de aptitudes mentales primarias (PMA), la batería de aptitudes mentales diferenciales (DAT), o el test de factor *g* de Cattell (Calero, 1997; Prieto y Pérez, 1993).

El presente estudio se ha desarrollado en las siguientes fases: (1) seleccionar una muestra de estudiantes de ESO en 3 centros de la Comunidad de Madrid; (2) formar grupos experimentales y de control en cada uno de los centros; (3) aplicar una serie de tests estandarizados a ambos grupos al comienzo del curso académico; (4) aplicar el programa de intervención durante ese curso académico (con el grupo experimental); (5) volver a aplicar una serie de tests estandarizados al finalizar el curso académico; (6) comprobar el grado de cambio en el rendimiento en los tests estandarizados, tanto del grupo experimental como del grupo de control.

A comienzos del curso se aplicaron los siguientes tests estandarizados en dos sesiones distintas: (a) *Test de Factor g de R.B. Cattell*. Es un test no verbal que exige deducir relaciones entre una serie de figuras. Se utilizó la escala 3, que incluye cuatro subtests: series, clasificación, matrices, y condiciones. El uso de este test en la valoración de programas de enseñar a pensar es bastante habitual. La fiabilidad  $r_{xx}$  de la escala 3 oscila entre 0.69 y 0.83. (b) *Test de Inteligencia General Factorial (IGF) de C. Yuste* (1997). Mide las aptitudes específicas espacial, abstracta, numérica y verbal. Permite estimar la inteligencia verbal, no verbal y general, como medidas sumario de las aptitudes específicas. Los índices de fiabilidad son: 0.697, 0.753, 0.850, 0.755, 0.845, 0.878, y 0.932, respectivamente. Las correlaciones de las escalas del IGF con el Otis oscilan entre 0.39 y 0.74. (c) *Test HSPQ de Personalidad de R. B. Cattell y M.D. Cattell* (1995). Mide una serie de rasgos pri-

marios de la personalidad y cuatro factores secundarios sobre los que se ha centrado este estudio: ansiedad, excitabilidad, extroversión e independencia. Los índices de fiabilidad (consistencia interna) de las escalas del HSPQ oscilan entre 0.22 y 0.41.

El programa Filosofía para Niños fue aplicado por tres profesores expertos durante el curso académico 95/96, siguiendo las pautas especificadas en la programación de actividades (Lipman et al., 1992). El programa se aplicó en las clases de Ética o de Aprender a pensar, dependiendo del centro, durante dos horas a la semana y durante todo el curso académico. La aplicación del programa supuso un total de 80 horas. Las pautas específicas de la aplicación del programa seguían la normativa reglamentada en los manuales del profesor de los subprogramas Harry y Lisa (Lipman, 1988, 1989). El grupo de control en cada uno de los centros estuvo compuesto por alumnos que no recibieron el programa FpN, sino clases de ética estándar. Al finalizar el curso académico (Mayo del 96) se volvieron a aplicar los tests comentados para comprobar el grado de cambio de ambos grupos.

### Resultados

Los resultados que se presentan son de dos tipos: (a) cambios de cada uno de los grupos (intra-grupo) considerando la fase de test y el retest, y (b) comparación de ambos grupos (inter-grupo) en la fase de test y en el retest. La Tabla 1 presenta los resultados en las escalas del IGF. Se observa que el rendimiento aumenta en el retest en todas las escalas y para ambos grupos. Las diferencias test-retest son estadísticamente significativas para los dos grupos.

Se calculó un análisis de varianza para averiguar si las diferencias entre los grupos resultaban estadísticamente significativas en la fase de test y en el retest para las diferentes medidas aportadas por el IGF. Ninguna de las comparaciones resultó significativa al 1%. Por tanto, antes de la aplicación del programa no existían diferencias de rendimiento en el IGF entre el grupo control y el grupo experimental. Tras la aplicación del programa, en la fase de retest, ambos grupos habían mejorado significativamente en todas las escalas de IGF con respecto a la fase de test, pero la diferencia entre los dos grupos no resultó estadísticamente significativa.

Tabla 1

Medidas del IGF en el Test y en el Re-Test, valores t, y valores p. N grupo control = 84. N grupo experimental = 91. Entre paréntesis se presentan los resultados del grupo experimental

MEDIDAS DEL IGF	GRUPO CONTROL (GRUPO EXPERIMENTAL)		RE-TEST MEDIA	D.T.	t	p.
	TEST MEDIA	D.T.				
Aptitud Espacial	55,6 (55,6)	23,3 (19,1)	63,3 (61,4)	21,3 (20,2)	-3,24 (-3,17)	.002 (.002)
Razonamiento Abstracto	52,54 (52,4)	20,2 (21,1)	60,8 (59,1)	20 (20,9)	-4,43 (-3,91)	.000 (.000)
Aptitud Verbal	54,8 (51,14)	18,9 (19,3)	59,9 (55,4)	20,1 (20,7)	-3,21 (-2,88)	.002 (.005)
Aptitud Numérica	63,6 (56,1)	21,3 (20,3)	68,5 (64,3)	20,7 (21,3)	-3,20 (-5,16)	.002 (.000)
Inteligencia Verbal	57,9 (51,9)	19,3 (18,9)	63,4 (58,5)	20,5 (20,7)	-4,53 (-5,13)	.000 (.000)
Inteligencia No-Verbal	52,7 (52,5)	21,7 (19,4)	61,7 (59,6)	19,9 (20,1)	-5,34 (-4,47)	.000 (.000)
Inteligencia General	54,9 (51,3)	19,9 (19,1)	62,7 (58,7)	20,2 (20,3)	-6,75 (-6,17)	.000 (.000)

Tabla 2

Medidas del Factor g en el Test y en el Re-Test, valores t, y valores p. N grupo control = 59. N grupo experimental = 56

Grupos	TEST		RE-TEST MEDIA	D.T.	t	p.
	MEDIA	D.T.				
Grupo control	103,5	18,3	106	16	-.86	.393
Grupo exp.	103,7	18,1	110	17,4	-2,04	.047

En la Tabla 2 se presentan los resultados en el Test de Factor *g* de Cattell. Se observa que ambos grupos aumentan en el retest. Sin embargo, el aumento del grupo experimental en el retest con respecto a la fase de test es estadísticamente significativo, mientras que el aumento del grupo de control no resulta estadísticamente significativo. El incremento de puntos en la escala estándar de CI es de 2.5 para el grupo de control y de 6.3 puntos para el grupo experimental. También se calculó un análisis de varianza para averiguar si las diferencias de rendimiento entre el grupo control y experimental en factor *g* eran estadísticamente significativas, tanto en la fase de test como en el retest. La diferencias entre los grupos en la fase de test no resultaron significativas ( $F=0,003$ ,  $p=0,956$ ) y tampoco resultaron significativas las diferencias entre ambos grupos en el retest ( $F=1,711$ ;  $p=0,194$ ). Con todo, debe observarse que las diferencias en el retest entre ambos grupos se aproximaron a la significación estadística, resultado congruente con la comparación intra-grupo. Por tanto, la mejora del grupo experimental en el retest con respecto a la fase de test es estadísticamente significativa, mientras que este no es el caso para el grupo de control. Sin embargo, las diferencias entre grupos no son significativas ni en el test ni en el retest. Hay que recordar que el test de Cattell constituye un estimador adecuado de la capacidad cognitiva general (*g*; véase Carroll, 1993) por lo que se podría afirmar que, en el presente estudio, la aplicación del programa FpN parece beneficiar la capacidad cognitiva general del alumno.

Finalmente, la Tabla 3 presenta los resultados obtenidos con el test de personalidad HSPQ. Ni para el grupo control ni para el grupo experimental se observan diferencias estadísticamente significativas en ninguna de las medidas. También se calculó un análisis de varianza para comparar ambos grupos en la fase de test y en el retest. Ninguna de las diferencias resultó estadísticamente significativa al 1%. Por tanto, no existen diferencias entre el grupo control y el experimental en la fase de test, y tampoco en el retest. Por consiguiente, la aplicación del programa no produce un cambio en las medidas de personalidad en el presente estudio.

### Conclusiones

Los resultados sugieren que el programa Filosofía para Niños produce un cierto impacto en el desarrollo de la capacidad general de los alumnos: la mejora del grupo experimental en el retest con respecto a la fase de test en la medida de factor *g* es estadísticamente significativa, mientras que la mejora del grupo control no lo es. Sin embargo, la aplicación del programa no parece producir una mejora diferencial en las aptitudes cognitivas más específicas y tampoco influye en una serie de rasgos básicos de la personalidad, a pesar de que el programa lo pretende explícitamente.

Según Nickerson et al. (1994) y el informe oficial de la Asociación Americana de Psicología sobre la inteligencia (Neisser et al., 1996) modificar las aptitudes cognitivas constituye una tarea compleja. Sin embargo, la presente investigación es un ejemplo más de que ello resulta posible, tal y como ha sido puesto de relieve en otros casos (Ceci y Williams, 1997; Colom, 1998; Hunt, 1995; Juan-Espinosa, 1997; Neisser et al., 1996).

En cualquier caso, debe tenerse en cuenta que la mejora se ha observado justo tras la aplicación del programa, por lo que antes de considerar concluyente el resultado alcanzado sería deseable realizar un seguimiento para comprobar si las mejoras se mantienen a largo plazo—seguimiento que suele resultar muy difícil. En el informe de la Asociación Americana de Psicología sobre la inteligencia se concluye lo siguiente respecto a los programas de mejora de las capacidades cognitivas: «una variable ambiental con una clara importancia es la presencia de una escolarización formal. Las escuelas afectan a la inteligencia de muchos modos, no sólo transmitiendo información específica, sino desarrollando ciertas habilidades intelectuales y actitudes. No ir a la escuela (o asistir a escuelas muy malas) tiene un claro efecto negativo sobre las puntuaciones en los tests de inteligencia. Los programas de pre-escolar e intervenciones similares suelen tener efectos positivos, pero en la mayor parte de los casos las ganancias desaparecen cuando termina el programa» (Neisser et al., 1996, p. 99 en la traducción de Andrés-Pueyo y Colom, 1998).

Es posible que un análisis aislado de los resultados del presente estudio lleve a la conclusión de que realmente no son tan optimistas. De hecho, se ha observado que las diferencias de rendimiento entre el grupo experimental y el de control en el retest, en la medida de la capacidad cognitiva general, no son estadísticamente significativas. Todo lo más, se ha comprobado que la mejora del grupo experimental es estadísticamente significativa en el retest con respecto a la fase de test, mientras que esa mejora del grupo control no es significativa. No obstante, veamos los resultados que se han encontrado en un contexto similar al empleado en el presente estudio, es decir, la Comunidad de Madrid, tras la aplicación del Programa Harvard, originalmente desarrollado para su aplicación en Venezuela (Herrnstein et al. 1986; véase Nickerson et al., 1994): el grupo control alcanza una puntuación de 31.19, mientras que el grupo experimental obtiene una puntuación de 32.85, algo más de punto y medio de diferencia (Megía, 1993). Por tanto, las ganancias del grupo experimental tras la aplicación del Programa Harvard en la Comunidad de Madrid, son sustancialmente menores que las obtenidas tras la aplicación del Programa FpN en el presente estudio.

El Programa Harvard no está dirigido a mejorar la capacidad general, sino que está estructurado en una serie de Unidades y Lecciones destinadas al desarrollo de algunas habilidades básicas

Tabla 3  
Medidas del HSPQ en el Test y en el Re-Test, valores *t*, y valores *p*. N grupo control = 83. N grupo experimental = 73. Entre paréntesis se presentan los resultados del grupo experimental

MEDIDAS DEL HSPQ	GRUPO CONTROL (GRUPO EXPERIMENTAL)		RE-TEST		<i>t</i>	<i>p</i> .
	TEST MEDIA	D.T	MEDIA	D.T		
Ansiedad	7.30 (7.30)	1.07 (1.11)	7.39 (7.30)	1.11 (1.09)	-.87 (.03)	.389 (.987)
Excitabilidad	5.49 (5.47)	1.07 (1.21)	5.29 (5.47)	1.12 (1.14)	.80 (.00)	.079 (1.000)
Extraversión	4.30 (4.43)	1.04 (1.21)	4.35 (4.42)	1.09 (1.12)	-.48 (.14)	.630 (.885)
Independencia	6.55 (6.41)	1.10 (1.14)	6.68 (6.41)	1.05 (1.15)	-1.34 (.01)	.185 (.992)

de pensamiento. Sin embargo, como criterio de valoración se eligió el Test de Factor *g* de Cattell. Ello redundaba en lo ya comentado: este tipo de medidas de carácter psicométrico se consideran habitualmente como criterio básico para constatar si el programa de mejora logra incrementar *g* o la capacidad general (Carroll, 1993, 1997; Colom, 1998; Jensen, 1998).

Hay buenas razones para que las mejoras en *g* se consideren como criterio de valoración de programas como el Harvard o Filosofía para Niños. Entre ellas se cuentan los datos considerados por la APA ya comentados (Neisser y col., 1996) respecto a la potencia predictiva de las medidas estandarizadas de inteligencia en el contexto escolar. En términos generales, resulta más probable que mejore el rendimiento escolar si mejoramos la capacidad general. Se supone que el incremento de *g* está llamado a incrementar los correlatos sociales de *g* (Colom y Andrés, 1999; Jensen, 1998). No obstante, ello puede ser una condición necesaria, pero en el momento presente no existen evidencias que puedan asegurar que será suficiente.

En suma, los resultados aquí presentados deben considerarse transitorios en tanto no se puedan confirmar con otras muestras y en tanto el seguimiento de la muestra original no demuestre que se evita la regresión a la media del grupo experimental. Este es el «caballo de batalla» de los programas de mejora hoy existentes. Como señala la APA (Neisser et al., 1996) las ganancias a largo plazo de los sujetos que reciben un programa de mejora (una vez termina su aplicación) están lejos de estar bien establecidas. En

cualquier caso, los resultados que aquí se han presentado pueden servir como instigadores de la realización de mayor número de estudios de este tipo en nuestro país que puedan ayudar a superar el panorama actual, no demasiado alentador, que transmiten habitualmente los expertos en psicometría. Aunque sería imprudente negar las evidencias disponibles que indican que mejorar la capacidad es significativamente complejo, una vez termina la aplicación de un programa, tal y como han señalado entre otros Carroll (1997), Gottfredson (1997), Herrnstein y Murray (1994), Hunt (1995), Lubinski y Humphreys (1997) o Neisser et al. (1996), sería también poco razonable concluir a partir de ahí que los programas de mejora no son tan prometedores como se suponía. Como han escrito Herrnstein y Murray (1994) al revisar este tipo de programas «son ilimitadas las posibilidades científicas que no se han probado todavía de mejorar ambientalmente la inteligencia. En principio, la inteligencia puede modificarse ambientalmente hasta límites desconocidos» (p. 390).

#### Agradecimientos

Agradecemos al Ministerio de Educación y Cultura (MEC) la beca (18.05.421B.441) de colaboración entre Centros de Enseñanza Secundaria y Departamentos universitarios concedida a los integrantes del equipo firmante del presente trabajo, gracias a la cual se pudo desarrollar la investigación.

#### Referencias

- Brody, N. (1992): *Intelligence*. San Diego: Academic Press.
- Brody, N. (1997): Intelligence, schooling, and society. *American Psychologist*, 52, 10, 1046-1050.
- Calero, M<sup>a</sup> D. (1997): *La mejora de la inteligencia*. Madrid: Pirámide.
- Carroll, J.B. (1993): *Human cognitive abilities*. Cambridge: Cambridge Univ. Press.
- Carroll, J.B. (1997): Psychometrics, intelligence, and public perception. *Intelligence*, 24, 1, 25-52.
- Cattell, R.B. (1990): *Manual del test de factor g de Cattell*. Madrid: TEA.
- Cattell, R.B. y M.D. Cattell (1995): *Manual del test de personalidad para adolescentes (HSPQ)*. Madrid: TEA.
- Ceci, S. y Williams, W. (1997): Schooling, intelligence, and income. *American Psychologist*, 52, 10, 1051-1058.
- Colom, R. (1997): Diferencias individuales en la educación. En M. Juan-Espinosa; R. Colom y M<sup>a</sup> Angeles Quiroga (coord.): *La práctica de la psicología diferencial en educación, clínica y deportes*. Madrid: Pirámide.
- Colom, R. (1998): *Psicología de las diferencias individuales. Teoría y Práctica*. Madrid: Pirámide.
- Colom, R. y Andrés, A. (1999): El estudio de la inteligencia humana: recapitulación ante el cambio de milenio. *Psicothema*, 11, 3, 453-476.
- Detterman, D. y Thompson, L.A. (1997): What is so special about special education? *American Psychologist*, 52, 10, 1082-1090.
- Ferrer Cervero, V. (1994): Algunas reflexiones sobre la evaluación y el proyecto Filosofía para Niños. *Aprender a pensar*, 9-10, 22-47.
- Gottfredson, L. (1997): Why *g* matters: the complexity of everyday life. *Intelligence*, 24, 1, 79-132.
- Herrnstein, R. y Murray, Ch. (1994): *The bell curve. Intelligence and class structure in american life*. New York: Free Press.
- Herrnstein, R.; Nickerson, R.; Sánchez, M. y Sweets, J.A. (1986): Teaching thinking skills. *Journal of American Psychological Association*, 41, 11, 1279-1289.
- Hunt, E. (1995): The role of intelligence in modern society. *American Scientist*. Traducción en A. Andrés Pueyo y R. Colom (coordinadores): *Ciencia y Política de la Inteligencia en la Sociedad Moderna*. Madrid: Biblioteca Nueva, 1998 (p. 257-286).
- Jensen, A. (1980): *Bias in mental testing*. London: Methuen.
- Jensen, A. (1998): *The g factor*. London: Praeger.
- Juan-Espinosa, M. (1997): *Geografía de la inteligencia humana. Las aptitudes cognitivas*. Madrid Pirámide.
- Lipman, M. (1988): *Lisa*. Madrid: Ediciones de la Torre.
- Lipman, M. (1989): *El descubrimiento de Harry*. Madrid: Ediciones de la Torre.
- Lipman, M. y col. (1992): *Investigación filosófica*. Madrid: Ediciones de la Torre.
- Lubinski, D. y Humphreys, L. (1997): Incorporating general intelligence into epidemiology and the social sciences. *Intelligence*, 24, 1, 159-201.
- Megía, M. (1993): Proyecto inteligencia «Harvard». Madrid: CEPE.
- Miranda Alonso, T. et al. (1990): Filosofía para Niños en 6<sup>o</sup> de EGB: investigación psicológica de los resultados obtenidos. *Aprender a pensar*, 1, 110-118.
- Neisser, U.; Boodoo, G.; Bouchard, T.; Boykin, A.; Brody, N.; Ceci, S.; Halpern, D.; Loehlin, J.; Perloff, R.; Sternberg, R. y Urbina, S. (1996): Intelligence: knowns and unknowns. *American Psychologist*, 51, 2, 77-101. Traducción en A. Andrés Pueyo y R. Colom (coordinadores): *Ciencia y Política de la Inteligencia en la Sociedad Moderna*. Madrid: Biblioteca Nueva, 1998. (p. 43-114).
- Nickerson, R.; Perkins, D. y Smith, E. (1994): *Enseñar a pensar. Aspectos de la aptitud intelectual*. Barcelona: Paidós.
- Prieto, M. y Pérez, L. (1993): *Programas para la mejora de la inteligencia*. Madrid: Síntesis.
- Schaie, W. (1996): *Intellectual development in adulthood. The Seattle longitudinal study*. Cambridge: Cambridge Univ. Press.
- Snow, R.E. (1989): Aptitude-Treatment Interaction as a Framework for Research on Individual Differences in Learning. En P.L. Ackerman; R.J. Sternberg y R. Glaser (Eds.): *Learning and Individual Differences: advances in theory and research*. Nueva York: Freeman.
- Snow, R. y Yalow, E. (1982): Educación e inteligencia. En R. Sternberg (Ed.): *Inteligencia humana, vol. III*. Barcelona: Paidós.
- Yuste, C. (1997): *Manual del test de inteligencia general factorial (IGF)*. Madrid: TEA.

