

Valoración del impacto de un programa de educación en valores en el último curso de secundaria obligatoria

Carlos Fierro-Hernández
Universidad de Málaga

Resumen:

A lo largo de cinco meses se llevó a cabo en alumnos de 4º de ESO de un instituto malagueño un programa de educación en actitudes y valores. Dos unidades docentes de dicho curso participaron en el programa y son considerados, en el análisis metodológico, como grupo (cuasi)experimental, mientras las otras tres unidades del mismo curso han servido como grupo de contraste o control. Los análisis han sido de comparación entre ambos grupos en cambio de actitudes (en un diseño test-retest), en rendimiento académico y en comportamiento en el aula. Los resultados no han sido en todo los esperados en el análisis test-retest. En otras comparaciones, sin embargo, los resultados en el grupo de intervención han sido mejores que los del grupo de contraste: en mejor rendimiento académico (notas de fin de curso) y en menor número de conductas de indisciplina. Además, en ambos grupos, y acaso como resultado indirecto del interés de los profesores en la educación en valores, se han observado mejoras significativas en las actitudes.

Palabras clave: educación en actitudes y valores, evaluación de programas educativos, Educación Secundaria, Andalucía.

Abstract: Assessment of the impact of a programme in values education in the last year of Compulsory Secondary Education

A programme of education in values and attitudes was implemented over five months with students from the fourth year of Secondary Education in a state establishment in Málaga. Two teaching samples from the above-mentioned course took part in the programme and are considered in the methodological analysis as quasi-experimental, whereas the other three samples of the same course have acted as a contrast/control group. The comparative analysis between both groups has been based on attitude change (in a test - retest model), academic performance and classroom behaviour. Results have not been fully as expected in the test-retest analysis. However, in other comparative cases, the results in the intervening group have been better than in the contrasting groups in terms of better academic performance (results at the end of the academic year) and of lower numbers of misconduct. Additionally, in both groups, and arguably as an indirect result of teachers' interest in the education in values, noteworthy attitude improvements have been perceived.

Key words: attitudes and values education, evaluation of education programmes, secondary education, Andalusia.

INTRODUCCIÓN

Una investigación que pretenda contribuir realmente al cambio y a la mejora en la enseñanza ha de articularse con la acción educativa real en un contexto también real. El presente informe se refiere a un estudio en esas condiciones con el objetivo de evaluar los resultados de un programa de educación en valores. El programa se desarrolló en un Instituto de Educación Secundaria con alumnos de 4º curso de la etapa obligatoria de secundaria (ESO). El estudio pertenece, por tanto, al género de la evaluación científica de programas: ha consistido en una investigación evaluativa con miras a orientar la práctica educativa. Su referente metodológico es el de la evaluación o valoración de programas (Stufflebeam y Shinkfield, 1987). Es un referente afín al de la investigación en la acción (Gimeno y Pérez-Gómez, 1992).

La investigación se realizó a instancias de un grupo de profesores de un instituto de la periferia de Málaga. El grupo había solicitado y obtenido de la Junta de Andalucía una subvención para llevar a cabo un estudio sobre las variables cruciales asociadas al comportamiento escolar en los alumnos adolescentes del mismo. De una primera fase de la investigación se ha informado ya en un primer trabajo publicado (Fierro-Hernández, 2002). Dicha fase tuvo carácter exploratorio, sin intervención propiamente dicha. En ella, en 178 alumnos del último ciclo de la Educación Secundaria Obligatoria, se estudiaron variables de contexto familiar, de personalidad y de capacidad intelectual, así como de comportamiento y de resultados académicos, examinando las relaciones entre ellas. Se trataba de obtener un perfil de la red de relaciones entre el entorno familiar y socioeconómico de los alumnos, sus aptitudes y rasgos personales y, en fin, sus resultados escolares.

De la primera fase del estudio se siguieron algunas conclusiones claras. El rendimiento académico y el comportamiento correcto en el instituto no se asociaban a variables sociales de nivel económico y cultural de las respectivas familias de los alumnos. Tampoco se asociaban a la capacidad intelectual de éstos, al menos a la capacidad que miden los tests al uso. La red de correlaciones observadas apuntaba, en cambio, a una clara relación con la dimensión básica de personalidad que Eysenck (1967/1970, 1981) conceptuó como psicoticismo e igualmente con actitudes de xenofobia, discriminatorias y agresivas. Por otro lado, el bajo rendimiento académico sí que se asociaba con el comportamiento inadecuado en el instituto. De los resultados mismos del primer estudio se desprendería una directriz educativa: la de fomentar actitudes opuestas a las referidas y que las contrarrestaran. Esta directriz, desde luego, converge con un objetivo educativo irrenunciable, el de la formación moral: educar en actitudes, en valores. La investigación podía y debía proseguir, en una segunda fase, como evaluación de un programa educativo en actitudes y valores. Cabía predecir que esta intervención redundaría en una mejora en la convivencia, en la disciplina escolar y en el propio rendimiento académico de los alumnos incluidos en el programa.

Sobre esas bases, y de acuerdo con el grupo de profesores interesados, se diseñó la segunda fase del estudio. La indagación empírica había de combinarse ahora con una intervención educativa en favor de determinadas actitudes. Ésta se realizaría en el contexto más amplio de un programa de educación en valores.

EL PROGRAMA EDUCATIVO

Se elaboró, pues, y se llevó a cabo un programa de educación en valores y actitudes. Para su puesta en práctica se escogió el último curso de la educación obligatoria (4º de ESO). Se eligió este curso y la correspondiente edad (en torno a los 16 años) por ser un curso y una edad que los profesores juzgan singularmente difíciles y a menudo conflictivos. Es elección que confiere especial significación educativa a los resultados que lleguen a obtenerse.

En el referido instituto, 4º de ESO se distribuye en cinco grupos de docencia. Con el fin de proceder en modo cuasi-experimental, se asignaron dos grupos a la condición del programa y los otros tres quedaron como grupos de contraste.

El programa educativo fue elaborado por el conjunto de profesores responsables del proyecto entero. Hacía especial hincapié en aspectos de actitudes y valores relativos a «los otros», a las personas de otras etnias o grupos, o nivel social, pero asimismo de otro sexo. Su realización estuvo a cargo de los profesores tutores en la correspondiente hora semanal de tutoría a lo largo de cinco meses, desde comienzos de enero hasta finales de mayo. Se consideró duración suficiente como para poder esperar resultados de la actuación. La intervención educativa llevada a cabo puede ser calificada de intensiva, específica y complementaria, aunque dentro del horario escolar.

El presente informe, sin embargo, se refiere no al contenido del programa educativo en las aulas, sino a la valoración de los resultados del mismo.

DISEÑO DE INVESTIGACIÓN EVALUATIVA

Entre los posibles modos de llevar a cabo la valoración del programa, se ha elegido un procedimiento que puede ser calificado de cuasi-experimental: asignación de un conjunto de sujetos a una condición de tratamiento, de intervención educativa complementaria y específica (grupo «experimental», suele decirse en abreviatura), mientras otro conjunto, sobre el que no se interviene, rendimiento y al queda como grupo «de control» o contraste. Es un diseño tradicional en investigación educativa recogido ya en la obra clásica de Campbell y Stanley (1966/1973). Los grupos D y E de 4º de ESO formaron el «grupo experimental», de intervención, y los otros tres grupos, A, B y C, sirvieron de «grupo de control».

Dentro de ese esquema básico cabían dos variantes de diseño. El más sencillo de ellos presume que el grupo experimental y el de control no difieren entre sí, salvo por azar: por una aleatoriedad estadísticamente mensurable y controlable. En esas

condiciones, basta con una sola medición o evaluación, la posterior, tras haber realizado el programa o intervención. Otro diseño algo más complejo y realmente preferible no realiza tal presunción y se aplica a una medición o evaluación en dos momentos: antes y después de la intervención o tratamiento (Fierro, 1987, para una descripción de ese paradigma metodológico). Este diseño, que implica exploración pre- y post- de los sujetos, es, por tanto, de «g test-retest».

El estudio se ha realizado en un diseño en dos momentos, test-retest, para la mayoría de las variables: todas las de comportamiento escolar y de actitudes. Para algunas variables, sin embargo, por su propia naturaleza, la doble medición no era posible (así, las notas de final de curso) o era difícil de justificar (el número de faltas disciplinarias). Para dichas variables, el estudio ha operado con evaluación única.

Sea en formato test-retest o en medición única, el presente estudio obedece a un enfoque de «g investigación en la acción», que también es, por otro lado, de «g investigación-evaluación-intervención». En tal enfoque, la investigación se realiza en la acción misma ordenada a transformar la realidad bajo estudio; se evalúan actitudes y comportamientos de los sujetos al propio tiempo que se interviene educativamente sobre ellos. Es un enfoque que responde al criterio, a la vez ético, técnico y científico de «ninguna evaluación sin intervención, ninguna intervención sin evaluación».

SUJETOS

El número total de sujetos evaluados ha sido de 142. De ellos, en los análisis, hubo que excluir a siete por distintas razones, sobre todo por carencia de datos para muchas variables. El número de sujetos bajo estudio queda, pues, en 135: de ellos, 53 en el grupo «experimental» y 82, en el de «control».

En algunas variables, las de doble evaluación, pre- y post-, el número ha quedado todavía más reducido. No se pudo computar a aquellos alumnos que faltaron a clase uno u otro de los dos días en que se aplicaron Escalas psicológicas: faltaba en ellos asimismo una evaluación. Para los correspondientes análisis quedan, pues, 92 sujetos: 49 en el grupo experimental y 43 en el de control. Este número se considera suficiente para un género de estudio, el cuasi-experimental, donde a menudo se trabaja incluso con grupos más pequeños.

INSTRUMENTOS, VARIABLES Y PROCEDIMIENTO

Las variables elegidas para el estudio de valoración del programa han sido aquéllas que en el estudio anterior (Fierro-Hernández, 2002) aparecieron relacionadas con rendimiento y buen clima de convivencia, mientras se presumen sensibles a una actuación educativa. No se han recogido variables sociales, familiares y de capacidad intelectual. La total ausencia de lazos fiables en aquel estudio hacía ocioso volver sobre ellas en una reiteración que sólo contribuiría a oscurecer los resultados globales.

La evaluación de variables correspondientes al diseño test-retest se realizó: la primera vez, previa, a mitad de enero; la segunda, posterior, a finales de mayo. Dos

han sido los instrumentos para esta evaluación: una *Hoja de registro de Incidencias y una Escala de Personalidad y Actitudes*. Cada uno de ellos incluía diversas variables.

La *Hoja de registro de Incidencias* (abreviada en adelante como HI) fue elaborada para la presente investigación de acuerdo con los profesores. Es una plantilla para observación y registro de conductas. En ella, los profesores han de registrar, día a día durante una semana, distintos comportamientos de agresión, daños o indisciplina observados en los alumnos. Son seis los apartados de incidencias indeseables que se computan y registran: 1) Agresiones a compañeros; 2) Daños materiales; 3) Conductas Desconsideradas; 4) Actos de Indisciplina; 5) Actos de Perturbación del orden en la convivencia, y 6) Suciedad.

Hubo dos semanas de cómputo y registro en HI: en la primera semana lectiva de enero y en la última semana de mayo. Los profesores de cada grupo docente rellenaron la *Hoja* en esos dos distintos momentos.

Una doble evaluación test-retest se efectuó asimismo mediante una *Escala de Personalidad y Actitudes* (en adelante, EPA), preparada para el caso. Sus ítems fueron extraídos de cuestionarios y escalas de uso frecuente en Psicología. En dos de sus dimensiones, conceptuadas como neuroticismo y psicoticismo, los ítems fueron extraídos del *Cuestionario de Personalidad para Jóvenes* (EPQ-J), de Eysenck y Eysenck (1982). Es sabido que neuroticismo constituye un factor de vulnerabilidad e inestabilidad emocional, mientras psicoticismo lo es de propensión a reacciones asociales o antisociales. En esta aplicación de EPA, neuroticismo y psicoticismo constaban, respectivamente, de 20 y de 17 ítems.

La *Escala* EPA evalúa asimismo actitudes socialmente valoradas, relativas a comportamientos discriminatorios y antisociales (o asociales). En el estudio anterior (Fierro-Hernández, 2002), las actitudes aparecieron asociadas con conductas inadecuadas en el aula y en el Instituto. En EPA, se evalúan, en concreto, actitudes de Hostilidad y Discriminación, incluyéndose algunos ítems relativos a discriminación por razón de sexo. Se trata de evaluar en los sujetos: 1) la predisposición al uso de la fuerza y la violencia en las relaciones sociales (Hostilidad, ocho ítems); 2) el racismo y la animadversión extra-grupo (Discriminación, 12 ítems).

La *Escala de Personalidad y Actitudes* se aplicó en un formato de cuádruple posible respuesta: sí / más bien sí / más bien no / no. Los alumnos la cumplimentaron en sus respectivas aulas, por grupos docentes, en dos ocasiones, en enero y en la primera semana de junio.

Otras variables merecedoras de consideración no podían ser medidas y analizadas -por su propia naturaleza- más que una sola vez. Es el caso del rendimiento académico en el que la calificación decisiva se produce a final de curso. Este rendimiento se ha computado según las calificaciones de los alumnos en junio, de acuerdo con una puntuación de 0 a 3, de suspenso a sobresaliente. Ciertos hechos objetivos de conducta escolar, los de un comportamiento en buena (o no tan buena) convivencia, tampoco encajaban en un diseño test-retest con medición diferencia-

da en enero y en junio. Su incorporación al estudio se ha hecho a través de conductas negativas concretas, poco frecuentes por fortuna, pero de fácil identificación y cómputo, procedentes de dos fuentes:

- Una primera fuente es la de los «partes» de faltas disciplinarias cometidas a lo largo del curso y de las que se dio traslado a la dirección del centro por los profesores al producirse el hecho denunciado. Esta variable, codificada como «Faltas», ha sido el número de ocasiones a lo largo del curso en que, de acuerdo con el régimen disciplinar de los institutos, conductas inapropiadas de un alumno fueron objeto de notificación a la dirección. En ellas, se han cuantificado con dos puntos los hechos más graves.
- La segunda han sido los «Envíos» de un alumno, por parte del profesor, al «aula de estudio» para apartarle del aula ordinaria cuyo normal desarrollo alteraba. Es una medida disciplinar practicada en el instituto y que encaja en la técnica conductista del «tiempo fuera». Se han computado también a lo largo de todo el año.
- De estas últimas variables –calificaciones, faltas disciplinarias, envío al aula de estudio– sólo se cuenta con una sola medida, un solo valor por tanto, el de las notas académicas de junio y, respectivamente, el número total de Faltas o Envíos a lo largo del curso, de octubre a junio. Se espera que también en estas variables aparezcan resultados atribuibles al programa educativo.

CUADRO I

Resumen de las variables estudiadas

EN DOBLE MEDICIÓN, DISEÑO TEST-RETEST

Hoja de registro de Incidencias (HI)

Agresiones

Daños

Desconsideraciones

Indisciplinas

Perturbaciones

Suciedad

Escala de Personalidad y Actitudes (EPA)

Neuroticismo

Psicoticismo

Hostilidad

Discriminación

EN UNA SOLA MEDICIÓN

Calificaciones
 Faltas
 Tiempos fuera

ANÁLISIS

Tres han sido los géneros de análisis de los datos, según el diseño, los contrastes y las mediciones -única o doble- efectuadas:

- En las variables para las que se cuenta con doble medición, las de HI y de EPA, se han obtenido puntuaciones diferenciales entre el primer momento y el segundo: entre la puntuación previa (enero) de cada sujeto y la posterior (mayo-junio) al programa educativo. Con ello, se pretendía ver si, en los grupos docentes (D y E) en los que se desarrolló el programa, las actitudes (EPA) y los comportamientos (HI) habían cambiado en la dirección prevista y deseada. Para contrastar esta hipótesis, se ha realizado un análisis de varianza comparando las puntuaciones diferenciales de «mejora» del grupo experimental con las del grupo de control (A, B y C).
- En el resto de las variables, las de una sola medición (calificaciones, faltas disciplinarias, envío al aula de estudio), se ha practicado asimismo el contraste mediante un análisis de la varianza, en este caso entre las puntuaciones absolutas de ambos grupos, experimental y de control.
- Se examinó, en fin, si con independencia del programa educativo realizado en el grupo experimental se habían dado cambios en todos los grupos docentes, cambios atribuibles entonces no al programa específico, sino al currículo y proceso educativo general en el instituto.

RESULTADOS

Se presentan los resultados del estudio según el orden de los correspondientes análisis que los generan.

DIFERENCIAS EN LOS CONTRASTES ANTES / DESPUÉS

La Tabla I recoge resultados del análisis de varianza en la comparación entre las medias de las puntuaciones diferenciales (diferencias pre/post) del grupo experimental (D y E, tomados conjuntamente) y el grupo de control (A, B y C, en conjunto). Realmente, en la mayoría de las variables el contraste en el análisis de varianza ha resultado no significativo. De las 13 variables del estudio, sólo en cuatro de ellas (todas de HI, ninguna de EPA) aparece significación estadística de interés. En aras de la simplicidad, la Tabla I presenta únicamente los valores promedio de puntua-

ciones diferenciales (todas ellas inferiores a la unidad) en las variables con contraste significativo. Se han omitido las otras nueve variables para no incluir en la Tabla resultados irrelevantes.

Las puntuaciones diferenciales se obtuvieron restando el valor post (en mayo-junio) del valor pre (en enero) de cada variable. Una puntuación diferencial de signo positivo quiere decir que la correspondiente variable cambió en la dirección deseada y prevista, la de disminuir la media de las incidencias de ese género en ese grupo, el experimental o el de control. El signo negativo quiere decir que incluso tales incidencias aumentaron.

Las cifras en las columnas correspondientes a los grupos experimental y de control son, todas ellas, inferiores a la unidad: se trata de promedios en las diferencias del primer momento al segundo por grupos de alumnos. Un valor positivo en esas celdas significa que la media en la correspondiente variable ha cambiado en la dirección de incrementarse todavía la media de incidencias en ese grupo. El valor negativo, por el contrario, significa que disminuyeron. La última columna expresa el nivel de probabilidad de error ($p <$) de realizar una inferencia errónea. Cuanto menor es ese valor, tanto más cierta es la inferencia. Desde luego, el bajo $p <$ permite afirmar que, para los grupos de referencia, la significación estadística es muy firme.

Como las variables todas de HI y, por tanto, las de la Tabla I, se refieren a comportamientos indeseables, las cifras positivas en esa Tabla proceden en la dirección educativamente deseada, mientras las negativas van en la dirección indeseable.

Los resultados de la Tabla I son pobres. En la mayoría de las variables, las ausentes de la Tabla, no han sido significativos; e incluso en las variables presentes son poco claros con respecto a hipótesis y predicciones previas. Según lo esperado, el grupo experimental, el del programa educativo, ha evolucionado mejor que el de control en perturbaciones y en agresiones; pero, en contra de ello, no ha evolucionado tan bien como el de control en las otras dos variables.

TABLA I
Contraste de las diferencias promedio pre / post

	Grupo control (N= 82)	Grupo experimental (N=53)	p <
Perturbación de la clase	-.43	.91	.001
Desconsideración	.26	.04	.002
Indisciplina	.29	-.21	.016
Agresiones	-.11	.11	.004

Los valores de las casillas son diferenciales: resultan de restar de los valores del momento «pre» los del momento «post». El signo negativo significa, por tanto, que hubo, en promedio, menos hechos de la correspondiente categoría en el momento posterior a la intervención.

DIFERENCIAS EN CALIFICACIONES Y EN DISCIPLINA

En este análisis, se supone que al inicio del programa educativo (enero) el grupo experimental (D, E) y el de control (A, B, C) no difieren más que por azar. Se analizan y contrastan, por tanto, los valores de las tres variables recogidas una sola vez (junio): las calificaciones de final de curso, el número total de Faltas (partes disciplinarias) y el de «tiempos fuera» (envíos al aula de estudio). Los resultados constan en la Tabla II. Las Faltas y los Tiempos fuera se miden en unidades de ocasiones. Las cifras en las calificaciones académicas corresponden a la puntuación obtenida a través de las diferentes materias del curso y en promedio de los respectivos alumnos de cada grupo: el suspenso es valor 0, mientras que 1 equivale a aprobado, 2 a notable y 3 a sobresaliente.

De la Tabla II, se desprende que los alumnos del grupo experimental tienen mejores valores promedio que el grupo de control en todas las variables. En calificaciones académicas, el grupo experimental supera en medio punto al grupo de control. En cambio, queda por debajo en la media de los otros comportamientos, que son precisamente los indeseados. Todos los resultados son significativos con $p < .001$, lo que permite un muy alto grado de confianza en las inferencias obtenidas.

TABLA II
Contraste de medias en calificaciones y en medidas disciplinarias

	Grupo control (N=82)	Grupo experimental (N=53)	p <
Envíos a aula estudio (tiempos fuera)	2.38	.41	.001
Partes de disciplina (faltas)	1.35	.19	.001
Calificaciones académicas	1.27	1.78	.001

CAMBIOS A LO LARGO DEL CURSO EN AMBOS GRUPOS

Al margen de la hipótesis principal del estudio sobre el impacto del programa, es plausible la conjetura de que en todo el instituto, o al menos en los profesores del curso 4° de ESO, se haya producido una especial insistencia educativa en ciertas actitudes y conductas, que son precisamente las evaluadas en esta investigación.

Para tratar de confirmar dicha conjetura, se han comparado los valores medios en todas las variables que sin duda han sido objeto de especial atención en este curso 4° de ESO, tanto dentro del programa educativo específico, cuanto en la docencia de las distintas materias a cargo de diferentes profesores.

Se ha realizado un contraste de medias, en análisis de prueba «t» para muestras relacionadas, comparando en todos los grupos (A, B, C, D, E) las puntuaciones en dos momentos, enero y mayo-junio, para las variables que parecen cruciales al res-

pecto: en lo comportamental, agresiones en HI; en actitudes, las de hostilidad, xenofobia y psicoticismo, de EPA. La Tabla III presenta los resultados.

TABLE III
Contraste «t» de medias en actitudes y agresiones

	Momento 1º (enero)	Momento 2º (mayo-junio)	p <
Hostilidad	9.66	8.86	.023
Xenofobia	16.31	14.30	.001
Psicoticismo	8.30	7.11	.012
Agresiones	.08	.06	n.s.

Puede observarse que todos los valores en estas variables han disminuido a lo largo de los trimestres segundo y tercero del curso. El descenso ha sido significativo en todas las variables de actitudes, pero no en agresiones, donde la variación no resulta significativa.

DISCUSIÓN Y CONCLUSIONES

El estudio había sido planteado como investigación en la acción. En consonancia con ello, atribuía gran importancia a la doble medición de algunas variables en un diseño test-retest. En ese planteamiento, los resultados de la Tabla I son estadísticamente significativos sólo para unas pocas variables y, además, lo son en direcciones opuestas. No debe sorprender que entre el grupo experimental y el de control no aparezcan diferencias significativas en factores de personalidad, como neuroticismo y psicoticismo. Era de esperar así, puesto que la correspondientes subescalas de EPA se refieren a factores personales no fácilmente cambiantes. Otros resultados contradicen las predicciones del estudio: el de que no haya habido cambios en las actitudes de xenofobia, discriminatorias y agresivas. Además, parecen invalidar la eficacia del programa educativo llevado a cabo con el grupo experimental.

Ni siquiera los resultados significativos de la Tabla I, relativos todos a HI, confirman las hipótesis del estudio. En efecto, el grupo experimental ha evolucionado mejor que el de control en perturbaciones y en agresiones, pero comparativamente peor en las otras dos variables. Es un resultado de difícil interpretación. La más plausible quizá sea que el programa de intervención educativa se ha centrado en actitudes de buena convivencia, no agresivas, más que en otros comportamientos: es, por tanto, comprensible que haya obtenido resultados ahí, no en otras incidencias. Aun entonces, sin embargo, no se entiende por qué ese resultado aparece en conductas y no en las correspondientes actitudes. El análisis test-retest, en suma, por sí solo, no permite confirmar las hipótesis del estudio, ni tampoco la confianza depositada en el programa educativo.

Distintos y más alentadores son los resultados que se desprenden al comparar el grupo experimental y el grupo de control en evaluación sólo posterior (Tabla II).

Muestran que la implementación de un programa específico de educación en valores se asocia a un mejor logro académico en los alumnos y también a un menor número en incidencias de relieve, tales como las faltas disciplinarias y los tiempos fuera por envío al aula de estudio. En estas variables, de gran importancia personal para los alumnos y también institucional para el clima social del instituto, sí que se confirman las hipótesis: las diferencias entre uno y otro grupo son sólidamente significativas; el programa educativo aparece claramente eficaz. El flanco más débil de esta conclusión es que no se deriva del diseño principal de análisis de los datos. Resultados semejantes, que se esperaban, no se han producido en el análisis test-retest.

Consideraciones parecidas a las derivadas de la Tabla II se siguen de la Tabla III. Sus datos se refieren a todos los grupos docentes de 4° de ESO en una comparación de los sujetos con ellos mismos en dos momentos distintos del curso: en enero y en mayo-junio, antes y después, por tanto, del programa educativo, aunque al margen de haberlo recibido o no. La Tabla III pone de manifiesto una mejora significativa en las variables de actitud (de EPA), en conductas de agresión registradas en aula e incluso también en un factor de personalidad, en psicoticismo. Son resultados muy positivos desde el punto de vista pedagógico.

Las mejoras recogidas en la Tabla III no pueden ser atribuidas de forma específica al programa educativo, que se llevó a cabo tan sólo en dos de los grupos docentes, no en todos ellos. Han de ser atribuidas a otras causas. Entre ellas, el mejor candidato es la escolarización misma: el hecho de que de enero a mayo transcurren cinco meses no de mero paso del tiempo, sino de trabajo educativo, haya habido, o no, entretanto un programa específico. El patrón de resultados de la Tabla III es por ello muy alentador con respecto no al programa educativo complementario, sino al trabajo educativo general realizado en el instituto. De forma clara, estadísticamente significativa, de enero a junio ha habido mejora en actitudes de los alumnos de 4° de ESO.

Desde el punto de vista metodológico, las conclusiones derivadas de las Tablas II y III habrían de ser sometidas a prueba de replicación en estudios de más larga duración dentro de unos mismos grupos: con análisis de un curso a otro y no sólo en distintos momentos de cada curso; con otros indicadores y variables. Sin embargo, en un enfoque pedagógico, de práctica educativa, dichas conclusiones son suficientemente firmes y confortadoras: tanto la docencia común en 4° de ESO como el programa específico del instituto a lo largo de dos trimestres han tenido efectos positivos por lo que toca a actitudes y a comportamientos de convivencia pacífica.

REFERENCIAS BIBLIOGRÁFICAS

- CAMPBELL, D. T.; STANLEY, J. C.: *Experimental and quasi-experimental designs for research / Diseños experimentales y cuasi-experimentales en la investigación social*. Chicago / Buenos Aires, Rand McNally / Amorrortu, 1966/1973.
- EYSENCK, H. J.: *The biological basis of personality / Fundamentos biológicos de la personalidad*. Springfield, Illinois / Barcelona, Charles, C., Thomas, Publisher / Fontanella., 1967/1970.
- (1981): *A model for personality*. Nueva York, Springer.
- EYSENCK, H. J.; EYSENCK G.: *Cuestionario de personalidad EPQ-J*. Madrid, Tea, 1982.
- FIERRO, A.: «El paradigma test / entrenamiento / retest en estrategias cognitivas de retrados mentales», en *Siglo Cero*, 109 (1987), pp. 26-43.
- FIERRO-HERNÁNDEZ, C.: «Patrón de rasgos personales y comportamiento escolar en jóvenes», en *Revista de Educación*, 329 (2002), pp. 373-392.
- GIMENO, J.; PÉREZ. GÓMEZ, A.: *Comprender y transformar la enseñanza*. Madrid, Morata, 1992.
- STUFFLEBEAM, D. L.; SHINKFIELD, A. J.: *Evaluación sistemática: guía teórica y práctica*. Madrid / Barcelona, MEC / Piados, 1987.