

REFLEXIVIDAD-IMPULSIVIDAD. VARIABLES RELEVANTES. ÚLTIMOS HALLAZGOS. PERDURABILIDAD DE LOS RESULTADOS OBTENIDOS MEDIANTE UN PROGRAMA DE INTERVENCIÓN

por BERNARDO GARGALLO LÓPEZ

*Escuela Universitaria de Formación del Profesorado de E.G.B.
Universidad de Valencia*

1. Introducción

A lo largo de todo un curso escolar (1986-87), realizamos una investigación en el ámbito de la Reflexividad-Impulsividad —estilo cognitivo que, conceptualizado por Kagan (Kagan, 1965, a, b y c, 1966, y 1987) como un constructo teórico bipolar que implica latencia, demora o tiempo previo a la emisión de la respuesta ante un problema con incertidumbre y precisión o exactitud en la elección hace ya más de 20 años, se ha convertido en un tema ya clásico en la Psicología y Ciencia de la Educación actuales— que pretendía, entre otros objetivos, incrementar la reflexividad de varios grupos de alumnos de 8.º de E.G.B. (Primaria). Para ello elaboramos un programa de intervención educativa y lo llevamos a término. El N de la muestra era de 201 sujetos, españoles, pertenecientes a cinco escuelas públicas diferentes, repartidos en 12 grupos, 6 experimentales y 6 de control, sobre los que actuaron 5 experimentadores distintos, uno de los cuales era el conductor de esta investigación.

Partimos de la constatación de que la reflexividad se ajusta más que la impulsividad a los requerimientos que la sociedad plantea al sujeto y de que, por tanto, le va a permitir adaptarse mejor a la misma, ser más

feliz y tener más éxito (Jelsma, O. y Peters, J.M., 1989; Schachar, R. y Logan, G.D., 1990 y Walczyk, J.J. y Hall, V.C., 1989). Situándonos en una perspectiva intervencionista planteamos el incremento de reflexividad como un objetivo educativo posible.

Además de valorar la *eficacia* del programa de intervención elaborado, en cuanto a incremento de reflexividad y a la *estabilidad* y *generalización* de los posibles resultados positivos obtenidos o a su *perdurabilidad*, pretendíamos estudiar una serie de *variables* fundamentales en el tema de la R-I, así como someter a crítica algunas de las afirmaciones de diversos autores, clásicos en la temática, con respecto a alguna de ellas.

En cuanto a la mejora de reflexividad hay datos de numerosos estudios previos que refieren un moderado éxito en ocasiones, un fracaso notable en otras y, en algunas, muy buenos resultados, con el uso de diversas técnicas y estrategias tales como *Demora forzada* (Gaines, 1971; Heider, 1971; Kagan, Pearson y Welch, 1966; Schwebel, 1966), *Reforzadores* (Briggs, 1968; Briggs y Weinberg, 1973; Errickson, Wyne y Routh, 1973; Heider, 1971; Loper, Hallahan y Mckinney, 1982; Scher, 1971; Weinberg, 1969), *Aprendizaje Observacional o Modelado* (Cohen y Przybycien, 1974; Debus, 1970 y 1976; Denney, 1972; Kagan, Pearson y Welch, 1966; Meichenbaum y Goodman, 1971; Ridberg, Parke y Hetherington, 1971; Yando y Kagan, 1968), *Enseñanza de estrategias cognitivas adecuadas de escudriñamiento* (Cow y Ward, 1980; Debus, 1976; Egeland, 1974; Heider, 1971; Meichenbaum, 1971 y 1977; Meichenbaum y Goodman, 1969 y 1971; Zelniker, Jeffrey, Ault y Parsons, 1972; Zelniker y Oppenheimer, 1974 y 1976), *Aprendizaje Autoinstruccional* (Cow y Ward, 1980; Meichenbaum, 1971, 1977 y 1981; Meichenbaum y Asarnow, 1979; Meichenbaum y Goodman, 1969 y 1971) o técnicas basadas en la teoría de orientación cognitiva (CO), tales como el *Plan Training* (Plan de entrenamiento o PT) (Kreitler y Kreitler, 1976 y 1982; Zakay, Bar-El y Kreitler, 1984).

De todas formas, es bastante general la constatación de la falta de perdurabilidad y estabilidad de resultados, así como la ausencia de generalización y transfer. Dicho con otras palabras: en algunas ocasiones se consigue un cierto éxito inmediatamente después de la aplicación de la técnica de que se trate, pero es bastante general la tendencia a la falta de perdurabilidad. Los sujetos de los experimentos, en muy poco tiempo, olvidan lo aprendido y retornan a sus estrategias habituales anteriores. Tienen, además grandes dificultades para transferir lo aprendido a otras situaciones similares (Egeland, 1974 y Meichenbaum, 1977).

En ocasiones se ha conseguido una cierta estabilidad y alguna evidencia de generalización de estrategias reflexivas 8 semanas después (Zakay, Bar-El y Kreitler, 1984) y hasta 4 y 8 meses más tarde (Debus, 1976), pero eso no es lo habitual.

Animados por estos hallazgos nos planteamos como un objetivo fundamental *el incremento de Reflexividad* en los sujetos sometidos a tratamiento experimental por efecto de la aplicación del programa de intervención que elaboramos y *la perdurabilidad, estabilidad y generalización* de los efectos positivos logrados con el mismo.

Tras un período de observación asistemática pasamos a otro en que ésta se realizó de modo más concienzudo y sistemático sobre los sujetos del experimento, observando en gran número de ellos características denotativas de impulsividad cognitiva: bajo rendimiento en exámenes, pruebas y trabajos, por el simple hecho de no haber leído detenidamente los enunciados de las preguntas o cuestiones, tendencia a responder con excesiva rapidez y sin tomarse el tiempo necesario, no agotar el tiempo concedido para pruebas o exámenes, falta de concentración y atención, falta de estrategias adecuadas de solución de problemas y, cuando éstas existían, falta de continuidad en su uso, etc...

2. Hipótesis

Partiendo de estas observaciones previas, de los datos de otras investigaciones disponibles, de los de nuestras propias investigaciones (Gargallo, 1985, 1987 y 1989) y de nuestra propia elaboración, planteamos la hipótesis de trabajo en los siguientes términos:

2.1. Pensamos que, partiendo de una situación de inexistencia de diferencia significativa de medias en R-I entre los grupos experimentales y los de control (dato que se comprobará en un Pretest), ésta se producirá por efecto de la aplicación del programa de intervención (para ello usaremos los datos de un 1.º Postest) y se mantendrá después de un tiempo intermedio prudencialmente largo, unos cuatro meses, de forma que los resultados obtenidos con el programa serán consistentes (tomaremos para ello los datos de un 2.º Postest).

Pensamos que esta diferencia se producirá en cada uno de los grupos experimentales inmediatamente después del tratamiento y que se mantendrá dejando transcurrir el mismo período antes reseñado y que esto no ocurrirá en los grupos de control, tomando datos de cada uno de los Postests y cotejándolos con el pase anterior del test.

2.2. Pretendíamos corroborar la inexistencia de diferencia significativa en la ejecución del MFF20, test operacionalizador de la R-I, entre los varones y las hembras (sexo). Con el MFFT de Kagan ha sido una tendencia generalizada la inexistencia de diferencias sexuales significativas al ejecutarlo (Lewis, 1971; Messer, 1976; Salkind, 1978; Ward, 1968; Zucker y Stricker, 1968). En la misma línea van los datos de Cairns y Cammock (Cairns y Cammock, 1984 y 1989) para el MFF20. Nosotros encontramos la misma tendencia en una investigación previa con el MFF20 efectuada entre 1984 y 1985 (Gargallo, 1985 y 1987).

2.3. Queríamos someter a crítica la afirmación comúnmente sostenida de efecto de techo para la R-I a los 9-10 años. Esta edad sería la mejor para la operacionalización de la R-I por el MFFT y, a partir de ella, ya no se seguiría progresando en reflexividad (es decir, utilizando mayores tiempos de latencia previa a la respuesta y cometiendo, con ello, menos errores) sino en eficacia para resolver la tarea (menos tiempo precisado para cometer menos errores). Esta es la línea de argumentación de Egeland y Weinberg (Egeland y Weinberg, 1976) y de Salkind y Nelson (Salkind y Nelson, 1980). Con los datos de nuestra anterior investigación ya nos cuestionamos estas afirmaciones, puesto que habíamos encontrado que la reflexividad seguía creciendo, por lo menos hasta los 13 años, pero como el N de la muestra era pequeño y el de esta edad también, no pudimos elaborar conclusiones firmes. Después dispusimos de los datos normativos de Cairns y Cammock (Cairns y Cammock, 1984) para el MFF20, que ofrecían puntuaciones muy similares de errores a los 10, 11 y 12 años, e incremento de latencia, lo que implica reflexividad parecida o ligeramente superior, pero los últimos datos, que dichos autores nos facilitaron amablemente, relativos a ejecuciones del MFF20 de diciembre de 1988, refieren una mayor impulsividad a los 11 y 12 años que a los 10: los sujetos emplean tiempos más breves y cometen más errores (Cairns y Cammock, 1989).

Queríamos, pues, si era posible, aclarar con nuestros datos estas posiciones tan diferentes y defendíamos que el estilo cognitivo sigue progresando hacia una mayor reflexividad (más latencias y menos errores) al menos durante las edades objeto de nuestro estudio (12-14 años edad).

2.4. Creíamos que la R-I tiene una moderada influencia sobre el Rendimiento Académico y apostábamos por una mejora de las calificaciones de los sujetos que incrementasen su reflexividad por efecto del tratamiento experimental, frente a los sujetos de control, en que esto no ocurriría.

2.5. Queríamos también analizar la estabilidad de las puntuaciones básicas de R-I: latencias, errores y puntuaciones I y E de Impulsividad y Eficiencia de Salkind y Wright (Salkind y Wright, 1977) obtenidas partiendo del MFF20, así como su validez de constructo como operacionalizador de la R-I, sin olvidar la posibilidad de dar respuesta a ciertas críticas (Gjerde, Block y Block, 1985), algunas de las cuales cuestionan todo el constructo (Block, 1987), sobre la ambigüedad y equívocidad de estas puntuaciones I y E.

Creíamos que el MFF20 era un instrumento más fiable, como medida y operacionalizador del constructo, más coherente y estable que el MFFT de Kagan, lo cual se reflejaría en una mayor estabilidad de las puntuaciones básicas de R-I (latencias, errores y puntuaciones I y E), con unos coeficientes de correlación más elevados que los obtenidos en otras investigaciones que han utilizado El MFFT.

3. Método

3.1. Sujetos

La muestra, con un N total de 201 sujetos, estaba constituida por 12 grupos, 6 experimentales y 6 de control, de cinco colegios públicos: 3 de la provincia de Valencia, 1 de la de Castellón y 1 de la de Teruel.

Había 107 chicos (53'2 %) y 94 chicas (46'8%), 26 eran de 12 años (13%), 147 de 13 (73%) y 28 de 14 (14%). Ocho de los doce grupos eran grupos «mosaico»: elegidos al azar al comienzo de la investigación para el grupo experimental y para el control, no constituían toda el aula sino sólo parte de los individuos de una clase. Esto se hizo para facilitar la labor de los colaboradores-experimentadores (cuatro profesores que aplicaron el programa, además del conductor de la investigación, para evitar al influencia indirecta del mismo sobre los sujetos del experimento partiendo de las expectativas que él podía tener. Por eso se aplicó el programa en cinco colegios diferentes y por cinco experimentadores distintos), ya que su trabajo era arduo pues, además de las sesiones de intervención, debían testar individualmente a cada uno de los sujetos del experimento de sus grupos en tres ocasiones (Pretest, 1.º Postest y 2.º Postest) y cada pase del test precisaba de alrededor de 20 minutos para su aplicación.

Los cuatro grupos restantes constituían todo el grupo de clase.

La clase social de procedencia era media y media-baja. El medio de procedencia era urbano y agrícola y el ámbito laboral de sus padres

barría prácticamente todo el espectro social. Todo ello favorece la posibilidad de generalización de resultados y la aplicación ulterior del programa a amplias capas de la población escolar.

3.2. *Instrumentos de medida*

Para operacionalizar la R-I utilizamos el MFF20 de Cairns y Cammock (Cairns y Cammock, 1978, 1984 y 1989). Es un test de emparejamiento perceptivo que consta de 20 ítems y otros 2 de prueba, cada uno de los cuales presenta simultáneamente al sujeto un dibujo modelo o estándar y seis variantes, casi idénticas del mismo, de las que sólo una es copia exacta. El sujeto tiene que averiguar cuál es. Se anota el tiempo antes de la primera respuesta para cada ítem y el número y orden de errores si se producen. Partiendo de la media de errores y de latencia del grupo, se clasificará al sujeto dentro de uno de los cuatro grupos clásicos: Lentos-Inexactos, Lentos-Exactos o Reflexivos, Rápidos-Inexactos o Impulsivos y Rápidos-Exactos. Nosotros escogimos el test de Cairns y Cammock y no el clásicamente utilizado de Kagan (Kagan, 1965 a, b, y c y 1966) porque los datos obtenidos partiendo del mismo revelan una mayor estabilidad y consistencia de puntuaciones de latencia y errores que los obtenidos partiendo del MFFT. Además las correlaciones que se generan partiendo de estas puntuaciones y de las puntuaciones I y E (de Impulsividad y Eficiencia) dan lugar a unos coeficientes más elevados que los obtenidos partiendo del MFFT, lo que se ha considerado clásicamente como una prueba de la validez del constructo.

3.3. *Temporalización*

La experiencia ocupó todo un curso escolar. Se tomó contacto con los otros experimentadores y se les entregó el material necesario en la 1.^a quincena de noviembre de 1986. El 1.^{er} pase del test MFF20 (Pretest) se efectuó en la 2.^a quincena del mismo mes. A continuación tuvieron lugar las 30 sesiones de intervención, durante los meses de diciembre, enero y febrero de 1987, a razón de tres semanales. El 2.^o pase del MFF20 (1.^o Posttest) se efectuó en la 1.^a quincena de marzo de 1987. El 3.^{er} pase del test (2.^o Posttest) se llevó a cabo de la 2.^a a la 4.^a semana de junio de 1987. Hay, pues, entre cada uno de los pases del test, un intervalo temporal que oscila de 3 meses y medio a 4 meses.

4. *El programa de intervención. Elaboración y aplicación*

Ya en 1984-85 habíamos desarrollado un programa que incidió sobre alumnos más jóvenes, de 5.º y 6.º de E.G.B., con notable éxito (Gargallo, 1985 y 1987). Nos faltaba corroborar la perdurabilidad de resultados y la influencia de la R-I sobre el Rendimiento Académico, y esto es lo que quisimos perfilar, junto a otros aspectos, con un programa que había crecido en cantidad y calidad, con una muestra mayor y con la colaboración de otros cuatro experimentadores. Ampliamos, para ello, el abanico de técnicas, basándonos en la bibliografía más reciente. Buscábamos incidir sobre los dos polos del constructo: tiempo conceptual o latencia y precisión-exactitud. Para ello seleccionamos las técnicas que habían mostrado un mayor nivel de eficacia en investigaciones anteriores y elaboramos un programa muy sencillo de aplicar y adaptado al nivel de los alumnos, que las aglutinara. Utilizamos:

1. Demora forzada.
2. Enseñanza de estrategias cognitivas adecuadas de escudriñamiento (scanning).
3. Aprendizaje Autoinstruccional. Enseñanza de estrategias de Autocontrol Verbal por el habla interna.
4. Plan Training (PT): Plan de entrenamiento en solución de problemas de respuesta abierta para incrementar la reflexividad, basado en la teoría de Orientación Cognitiva (CO).
5. Modelado participativo.
6. Reforzadores placenteros y aversivos. Utilizamos sistemáticamente el coste de respuesta. (Para un conocimiento más detallado de estas técnicas vid. Castillejo y Gargallo, 1989).

El programa, que constaba de 30 sesiones de alrededor de 20 minutos cada una, nada complicadas, perfectamente asimilables dentro del horario escolar y de la dinámica de la clase, con alto nivel de motivación y atractivo para los niños, se basaba en estas técnicas, tenía unos objetivos muy claros y unas estrategias muy perfiladas. Era muy funcional y su grado de eficacia se corroboraría en los dos Postests y en el subsiguiente análisis de los datos estadísticos recogidos.

5. *Resultados*

Los resultados obtenidos por los sujetos en cada pase del test en

latencia y errores se han utilizado para clasificarlos en los cuatro grupos clásicos. Partiendo de la media de cada grupo se puede hacer ya una primera valoración aproximativa en cuanto a la evolución del grupo en reflexividad, pero, como ello no es suficiente, el tratamiento de los datos se ha orientado en las siguientes direcciones fundamentales, de acuerdo con los puntos de la hipótesis:

5.1. Para valorar la *eficacia* del programa y la *perdurabilidad* de resultados llevamos a cabo diversos Análisis Discriminantes del tipo «Stepwise», centrándonos especialmente en los Anovas, Ancovas y Correlaciones Canónicas de los mismos, con dos introducciones diferentes de datos:

5.1.1. Partiendo de la 1.^a de ellas efectuamos diversos discriminantes Intragrupo [28], para confirmar si en cada uno de los grupos (los 12 grupos de la experiencia, más otro macrogrupo formado por los sujetos experimentales y otro formado por todos los de control) se establecía diferencia significativa de medias entre los resultados obtenidos en el 1.^{er} pase del test frente a los obtenidos en el 2.^o (Prestest-1.^o Postest) y también entre los resultados obtenidos en el 2.^o pase del test frente a los logrados en el 3.^o (1.^o Postest-2.^o Postest) y, si esta diferencia, de producirse, se debía al tratamiento experimental o si era algo espontáneo, de motu proprio, al margen de la intervención y por una evolución autónoma generada por la propia dinámica del grupo. Se comprobaría así si la intervención producía diferencias significativas entre los grupos y se éstas, una vez producidas, se mantenían después. Se efectuaron también pruebas T de diferencia de significación de medias para muestras pequeñas cuando el N de la muestra era reducido y había dudas sobre la validez de los datos aportados por el programa estándar de análisis de datos.

1.^o-2.^o pases del MFF20 (Datos del Prestest-Datos del 1.^o Postest)

...Sujetos experimentales:

Habíamos previsto un aumento significativo de reflexividad en los sujetos sometidos al programa de intervención (emplearían más tiempo y cometerían menos errores) por efecto del tratamiento y, cuando consideramos juntos los resultados del macrogrupo formado por *todos* los sujetos de grupos experimentales (N= 98), la hipótesis se cumplió en grado altamente significativo: menor número de errores y latencias más elevadas (Tabla 1). Grupo a grupo los resultados fueron similares: hubo incremento de reflexividad en un grado significativo en 5 de los 6 grupos experimentales (0'01) y en el 6.^o se dio aumento de eficacia (menos tiempo y menos errores).

...Sujetos de control:

Con los sujetos de control nos equivocamos en las previsiones: habíamos predicho niveles similares de R-I (parecido número de errores y tiempos de latencia semejantes del Pretest al 1.º Postest). Al analizar los resultados obtenidos por el macrogrupo formado por *todos* los sujetos de grupos de control (N= 103) se dio diferencia significativa de medias en errores, que bajaron, y no en latencias, que se mantuvieron muy similares con un ligero descenso: los sujetos, de motu propio, se habían hecho más eficaces para resolver la tarea (Tabla 2). Grupo a grupo los resultados fueron como siguen: de los seis grupos de control tres disminuyeron significativamente su número de errores y tres no lo hicieron: de estos últimos sólo uno los aumentó y ninguno de los seis incrementó significativamente su latencia; al contrario, la disminuyeron. Creemos que esta mejora en la eficacia para resolver la tarea se debe a la experiencia previa, al autoaprendizaje y a una elevada motivación de los sujetos de control para mejorar, sin haber recibido instrucciones para ello.

2.º-3.º pases del MFF20 (Datos del 1.º Postest-Datos del 2.º Postest)

Predijimos que las mejoras obtenidas por los sujetos experimentales corroboradas al comparar los datos del 1.º Postest y los del Pretest, se mantendrían en el 2.º Postest y también que los sujetos de control mantendrían niveles similares de R-I.

...Sujetos experimentales:

Al analizar los resultados obtenidos por todos los sujetos experimentales juntos (N= 98) se halló diferencia significativa de medias en media de errores: éstos seguían bajando, y no en media de latencias: los sujetos se habían hecho más eficaces resolviendo la tarea con muchos menos errores (del 1.º Postest al 2.º Postest) en un tiempo muy poco menor. Los resultados superaron las previsiones: más que estabilidad de resultados se dio una mejora de los mismos. Al analizar, por separado, los resultados obtenidos por los seis grupos experimentales, los resultados fueron muy similares.

Hay que hablar, pues, de *perdurabilidad y estabilidad* de resultados positivos del programa casi cuatro meses después de su aplicación. El tratamiento había sido, pues, sumamente eficaz (tabla 1).

...Sujetos de control:

Analizando los resultados obtenidos por todos los sujetos de grupos de control juntos (N= 103) no se dio diferencia significativa de medias ni

en errores ni en latencias: hay, pues, una gran estabilidad de resultados y un estancamiento de los mismos; los errores apenas si disminuyen y las latencias lo hacen en muy poco grado. Los resultados de R-I son muy similares. Lo mismo ha ocurrido al analizar los resultados de cada uno de los seis grupos de control por separado (tabla 2).

Se han cumplido, pues, las previsiones del primer apartado de la hipótesis en línea de *estabilidad, perdurabilidad y consistencia* de resultados del 1.º Posttest al 2.º (2.º-3.º pases del MFF20).

5.1.2. Partiendo de una 2.ª introducción de datos que diese pie a valorar un mayor número de variables se efectuaron otros análisis discriminantes [21] confrontando los datos obtenidos por los sujetos experimentales y los de control de cada colegio en cada uno de los 3 pases del test (Datos de experimentales en el Pretest-Datos de sujetos de control en el Pretest, Datos de sujetos experimentales en el 1.º Posttest-Datos de sujetos de control el 1.º Posttest, etc...) para calibrar si, ya de partida, existía diferencia significativa de medias en latencia y errores (creíamos que eso no iba a suceder) o si ésta se producía por efecto del tratamiento y se mantenía después de la intervención unos cuatro meses después (es lo que pensábamos). Esto se llevó a cabo en cada uno de los cinco colegios que participaron en la experiencia y también tomando los datos de *todos* los sujetos experimentales de la muestra frente a *todos* los de control porque el N de la muestra total (201 sujetos) era lo suficientemente importante como para ofrecernos elevadas garantías de fiabilidad, ya que en cuatro de los cinco colegios elegidos los grupos eran muy reducidos (su N oscilaba de 10 a 15 sujetos).

Pretest:

Tal y como se predijo, al considerar los datos aportados por *todos* los sujetos experimentales y por *todos* los de control, no se encontró diferencia significativa de medias ni en errores ni en latencias (Tabla 3).

Tampoco se dio en 4 de los colegios (en el colegio en que sí se dio debióse a que el experimentador del mismo malinterpretó nuestras instrucciones y realizó la adscripción de los sujetos a grupo experimental y de control después del Pretest y no antes y aleatoriamente, como estaba estipulado, sino precisamente en función de las puntuaciones obtenidas en el mismo, adscribiendo al grupo experimental a los sujetos más impulsivos y al de control a los más reflexivos. Le pareció lógico aplicar el programa de incremento de reflexividad a los más impulsivos, una buena lógica, aunque no era la nuestra para la experiencia. Así y todo no desestimamos sus datos que podíamos interpretar con esta salvedad).

1.º Posttest:

Al considerar *todos* los sujetos se halló diferencia significativa de medias en errores y latencias: los sujetos experimentales se habían hecho más reflexivos que los de control debido al tratamiento: cometían muchos menos errores y empleaban más largas latencias. Colegio a colegio se cumplió no en todos, pero sí en el de N mayor, que superaba a los 100 alumnos, con cuatro grupos de clase que constituían todo el alumnado de esas cuatro clases (Tabla 3).

2.º Posttest:

Al considerar el macrogrupo formado por todos los sujetos se halló de nuevo diferencia significativa de medias en latencias y errores: los experimentales seguían siendo más reflexivos que los de control (Tabla 3). Colegio a colegio se cumplió en unos en línea de mayor reflexividad, en otros en línea de mayor eficacia y en otros no se dio lo esperado. De nuevo el colegio de N mayor se ajustó en sus resultados exactamente a lo previsto.

Hay que constatar, pues, la eficacia del programa en cuanto a mejora de reflexividad en los sujetos experimentales frente a los de control inmediatamente después del tratamiento experimental (1.º Posttest) e igualmente en cuanto a perdurabilidad y estabilidad de resultados, ya que cuatro meses después los sujetos experimentales siguen siendo más reflexivos que los de control empleando más largas latencias y cometiendo menos errores. Queremos insistir en ello porque no es un dato usual en el contexto de las investigaciones sobre el tema. Los buenos resultados, su *consistencia, perdurabilidad y estabilidad*, la

TABLA 1
TODOS LOS EXPERIMENTALES EN PRETEST-1.º POSTEST

	X Pret.	Des. Típ.	X 1.º Pos.	Des. Típ.	GL	F(Anova)	p<(sig.)
Errores	7'09184	(5'42633)	1'81633	(2'59796)	1 y 194	75'36	p<0'01
Tiempo	26'31327	(11'8177)	33'07143	(12'0117)	1 y 194	15'76	p<0'01

TODOS LOS EXPERIMENTALES EN 1.º POSTEST-2.º POSTEST

	\bar{X} 1.º Pos.	Des. Típ.	\bar{X} 2.º Post.	Des. Típ.	GL	F(Anova)	p<(sig.)
Errores	1'81633	(2'59796)	1'17347	(1'55997)	1 y 194	4'41	p<0'05
Tiempo	33'07143	(12'0117)	29'94286	(10'9995)	1 y 194	3'62	No sig.

TABLA 2
TODOS LOS DE CONTROL EN PRETEST-1.º POSTEST

	\bar{X} Pret.	Des. Típ.	\bar{X} 1.º Pos.	Des. Típ.	GL	F(Anova)	p<(sig.)
Errores	5'87379	(4'94217)	3'88350	(3'67103)	1 y 204	10'76	p<0'01
Tiempo	26'55049	(9'97311)	24'2165	(9'05529)	1 y 204	3'09	No sig.

TODOS LOS DE CONTROL EN 1.º POSTEST-2.º POSTEST

	\bar{X} 1.º Pos.	Des. Típ.	\bar{X} 2.º Pos.	Des. Típ.	GL	F(Anova)	p<(sig.)
Errores	3'8835	(3'67103)	3'51456	(3'55314)	1 y 204	0'54	No sig.
Tiempo	24'2165	(9'05529)	23'79223	(10'0761)	1 y 204	0'10	No sig.

TABLA 3
TODOS LOS SUJETOS EXPERIMENTALES FRENTE A TODOS LOS DE CONTROL

	\bar{X} Exp.	Des. Típ.	\bar{X} Cont.	Des. Típ.	GL	F(Anova)	p<(sig.)
PRETEST							
Errores	7'09184	(5'42633)	5'87379	(4'94217)	1 y 199	2'77	No sig.
Tiempo	26'31327	(11'8177)	26'55049	(9'97311)	1 y 199	0'02	No sig.
1.º POSTEST							
Errores	1'81633	(2'59796)	3'88350	(3'67103)	1 y 199	21'04	p<0'01
Tiempo	33'07143	(12'0117)	24'2165	(9'05529)	1 y 199	35'05	p<0'01
2.º POSTEST							
Errores	1'17347	(1'55957)	3'51456	(3'55314)	1 y 199	35'94	p<0'01
Tiempo	29'94786	(10'9995)	23'79223	(10'0761)	1 y 199	27'2	p<0'01

amplitud de la muestra y la clase social de origen permiten apostar por la *generalizabilidad* de los efectos positivos al aplicar el programa a otros grupos similares y, por el tipo de muestra, estos grupos serán amplias capas de la población escolar.

Con ello queda confirmado el punto 1.º de la hipótesis.

5.2. Sexo

Tomando las puntuaciones obtenidas en el Pretest, antes de la intervención, por los varones (N = 107) y las hembras (N = 94) no hallamos diferencia significativa de medias ni en errores ni en latencias en Anova y Ancova, extremo este que concuerda totalmente con lo encontrado habitualmente al analizar esta variable en el contexto de las investigaciones de R-I.

	<u>Varones</u>		<u>Hembras</u>		GL	F(Anova)	Signif.
	\bar{X}	Des. Típ.	\bar{X}	Des. Típ.			
Errores	6'60748	(5'30851)	6'30851	(5'11592)	1 y 199	0'16	No sign.
Tiempo	26'4654	26'46542	26'43830	(11'47912)	1 y 199	0'00	No sign.

Se confirma, pues, el segundo apartado de la hipótesis.

5.3. Edad

Efectuamos un análisis discriminante del tipo stepwise tomando las puntuaciones obtenidas en el Pretest, antes de la intervención, ya que ésta podría distorsionar los resultados, de los 3 grupos de edad presentes en nuestra experiencia, 12, 13 y 14 años. Los resultados fueron:

	<u>12 años</u>		<u>13 años</u>		<u>14 años</u>		GL	F(Anova)	F(Ancova)	Sign.
	\bar{X}	Des. Típ.	\bar{X}	Des. Típ.	\bar{X}	Des. Típ.				
Errores	7'8	(2'94)	5'95	(4'92)	9'32	(6'32)	2 y 198	5'42	7'67	0'01
Tiempo	22'64	(6'61)	27'08	(10'74)	23'32	(11'88)	2 y 198	1'76		No sig.

Hemos constatado, pues, diferencias significativas en medias de errores, que no han existido en medias de latencia.

Nos parece interesante corroborar que el efecto de techo defendido por Salkind y Nelson (Salkind y Nelson, 1980), ya comentado previa-

mente, no se da para la edad de 9-10 años (10-12 según otros autores), como apuntaban, lo cual confirma los datos de nuestra primera investigación sobre R-I (Gargallo, 1985 y 1987), sino que la reflexividad sigue creciendo con la edad, al menos hasta los 13 años: de los 12 a los 13 años aumentan las latencias y disminuyen los errores. A los 14 años, y es un dato concordante con los de nuestra primera investigación, se produce un incremento en la media de errores y un descenso en la de latencia: los sujetos se hacen más impulsivos, dato que tampoco es concorde con la afirmación de Salkind de incremento de eficacia para resolver la tarea que, según él, se da con la edad: menos tiempo y menos errores, o errores muy similares, ya que aquí la media de errores es notablemente superior.

Nosotros, al igual que Salkind, encontramos en nuestra primera investigación más errores a los 12 años que a los 11 y menos latencias, pero a los 13 hallamos menos errores y más latencias, al igual que en esta última investigación; mayor reflexividad, en definitiva.

Estos datos, referentes a los 13 años, no los aporta ningún autor de los consultados por nosotros y esta edad no viene incluida ni en los baremos de Salkind ni en los datos normativos de Cairns y Cammock, y es importante por el incremento de reflexividad que se produce en la misma, al menos en nuestras dos investigaciones con diferentes grupos y muestras, y no por el incremento de eficacia que podría haberse esperado.

No queremos cerrar este apartado sin establecer antes una sencilla comparación entre las medias de edad en errores y latencias obtenidas en nuestra muestra y las que provienen de los baremos de Salkind (Salkind, 1978) y de Cairns y Cammock (Cairns y Cammock, 1984 y 1989). Los primeros son de niños estadounidenses y los segundos de niños norirlandeses. Únicamente podremos tomar los datos referentes a los sujetos de 12 años ya que, de las 3 edades con las que trabajamos, es de la única que refieren datos los autores antes aludidos, ninguno de los cuales llega a los 13 años.

Niños de 12 años

	Nuestros datos Medias	Salkind (1978) Medias	Cairns y Cammock (1984) y (1988) Medias	Medias
Errores	7'8	8'02	19'23	18'73
Tiempo	22'64	12'53	11'88	11'76

De nuestra investigación se desprende que los sujetos de nuestra muestra son más reflexivos que los de la de Salkind: emplean más tiempo en resolver la prueba y cometen menos errores. Si las puntuaciones de error no presentan gran diferencia ello se debe a que Salkind trabajó en el MFFT de 12 ítems y nosotros con el MFF20 de 20 ítems y, obviamente, se cometen menos errores en una prueba de 12 ítems que en otra similar de 20. Las puntuaciones de latencia sí presentan notable diferencia y ello no es afectado por la longitud de la prueba, ya que la media de latencia se saca dividiendo el tiempo total empleado en resolverla por el número de ítems de la misma, mientras que la puntuación de errores sale de errores totales cometidos por el individuo al responderla.

También son más reflexivos que los de la muestra de Cairns y Cammock y de una manera sustancial: emplean casi el doble de tiempo en resolver la tarea y cometen menos de la mitad de errores. En este caso la prueba empleada fue la misma: el MFF20 de dichos autores, con lo que la interpretación de resultados no ofrece ninguna duda.

Ya es más arriesgado extrapolar estos datos al conjunto de la población española, ya que la muestra era reducida, casi toda de la Comunidad Valenciana, y no puede representarla en su conjunto. Es ésta una tarea para ulteriores investigaciones y para trabajo en equipo.

5.4. R-I y Rendimiento Académico

Pretendíamos en un primer momento, valorar la relación existente entre la R-I y el Rendimiento Académico, expresado en las calificaciones de los sujetos. Lo hicimos tomando puntuaciones de latencia y errores, clásicas en el constructo, y también puntuaciones continuas «I» y «E» de Impulsividad y Eficiencia.

...Correlaciones de latencia y errores con calificaciones:

Al correlacionar errores con calificaciones encontramos coeficientes de correlación que oscilaban de $-0'30$ a $-0'35$, todos negativos, lo cual implica que a mayor número de errores en el MFF20 se obtienen peores calificaciones y viceversa. Los sujetos con más errores tienen peor rendimiento académico. Los coeficientes dan valores bajos, lo cual supone que la relación, aún estando presente, es muy moderada.

Al correlacionar latencias con calificaciones, los coeficientes obtenidos fueron siempre positivos, lo que implicaría que a mayor tiempo de reflexión corresponderían mejores calificaciones y mejor rendimiento, pero sus valores, que oscilan de $0'10$ a $0'16$, son despreciables.

...Correlaciones de puntuaciones I y E con calificaciones:

Al correlacionar estas puntuaciones con las notas se encontraron coeficientes de correlación que oscilaban de $-0'24$ a $-0'38$. Los valores obtenidos fueron siempre negativos, lo que implica que una puntuación positiva elevada de «I» (indicativa de Impulsividad) conlleva peores calificaciones y viceversa, y que una puntuación elevada de «E» (indicativa de Ineficiencia) implica peores calificaciones y al contrario. Los sujetos más impulsivos tienen peores notas académicas al igual que los más ineficientes; por contra, los reflexivos y eficientes consiguen mejores calificaciones que reflejan un rendimiento académico superior. No hay que olvidar, sin embargo, que los valores de los coeficientes son bajos, lo cual supone que existe relación o asociación pero que ésta es muy moderada.

..Análisis de Regresión Múltiple:

De acuerdo con las sugerencias y puntualizaciones de diversos autores (Ault, Mitchell y Hartman, 1976; Bentler y McClain, 1976; Lösel, 1980; Salkind y Wright, 1977), que preconizaban, junto a otras medidas, mantener continuas las puntuaciones de latencia y errores sin dicotomizarlas artificialmente por el procedimiento de división por la media y utilizar el Análisis de Regresión Múltiple, empleamos este procedimiento para predecir el Rendimiento Académico de los sujetos de la experiencia partiendo de las puntuaciones de latencia y errores de los mismos en los tres pases del test y efectuando 4 Análisis de Regresión Múltiple para pronosticar las calificaciones de Lengua y Matemáticas en junio del 86 y en junio del 87. Como variables predictoras se introdujeron las puntuaciones medias de latencia y errores y como variable criterio o predictando las medias de esas calificaciones.

Los coeficientes de correlación múltiple oscilan de $0'30$ a $0'35$ para los errores como predictores del rendimiento académico y desde $0'34$ a $0'39$ para las latencias. Los coeficientes de determinación (R^2) oscilan de $0'0957$ a $0'1296$ para errores y de $0'1175$ a $0'1584$ para la asociación de errores con latencias. Por eso los errores explican, en el punto más bajo, el $9'57\%$ de la varianza de las calificaciones de los alumnos, y el $12'96\%$ en el caso más alto. Las latencias sólo añaden al componente de variabilidad explicado por los errores un $2'18\%$ en el primer caso y un $2'88\%$ en el segundo.

Hay, pues, una ligera influencia-asociación entre errores cometidos al resolver el MFF20 y calificaciones que obtienen los sujetos y no hay asociación o relación entre latencias o tiempo empleado en resolver la tarea y Rendimiento Académico, a no ser que esas latencias se asocien con los errores, en cuyo caso su peso sigue siendo despreciable.

Todo ello es congruente con la línea argumental general derivada de los datos muy abundantes previos a este estudio que constatan que la influencia de la R-I sobre el rendimiento académico se da, pero en grado bajo. También lo es con los autores que afirman que es el componente de error y no el de latencia el que tiene más incidencia sobre el rendimiento (Block et al., 1974 y 1975; Block et al., 1986; Gjerde et al., 1985; Haskins y Mckinney, 1976; Nagie y Thwaite, 1979).

...Después de constatar esta influencia pasamos a delimitar, con la realización de diversos *Análisis de Varianza*, cómo se plasmaba, en concreto, con los datos de los sujetos de nuestra investigación:

Un 1.º Anova, al confrontar las calificaciones de los sujetos que, en el Pretest, habían resultado impulsivos, con las de los reflexivos en el mismo, encontró diferencia significativa de medias $p < 0'01$, resultado que confirmaba la influencia de la R-I sobre el Rendimiento Académico de los sujetos de nuestra investigación. Los reflexivos obtenían mejores calificaciones que los impulsivos en Lengua y Matemáticas en grado significativo.

Otros 8 Anovas, que pretendían corroborar si, como habíamos supuesto y formulado en nuestra hipótesis, el incremento de reflexividad en los sujetos sometidos al programa de intervención conllevaba un incremento significativo en sus calificaciones, cosa que no ocurriría en los grupos de control, dieron resultados contrarios a los esperados para los experimentales: ni los sujetos experimentales incrementaban significativamente sus calificaciones al mejorar en reflexividad ni al hacerse más impulsivos las disminuían significativamente: todos los experimentales, se hicieran más reflexivos o no, mejoraban sus calificaciones, pero no de un modo significativo. En los grupos de control había quienes mejoraban sus calificaciones y quienes las empeoraban, pero no de modo significativo ni tampoco por hacerse más reflexivos o impulsivos.

Se efectuaron otros 8 Anovas para determinar ahora si los sujetos que se habían hecho más Eficientes, en puntuaciones lineales «E», mejoraban sustancialmente sus calificaciones por efecto del programa de intervención y los resultados conseguidos fueron similares a los del caso anterior. En ningún caso se produjo diferencia significativa de medias entre las calificaciones que obtenían los sujetos que se volvían más eficientes en el MFF20 ni tampoco en los que se convertían en más ineficientes, y ello tanto en sujetos experimentales como en sujetos de control.

Otros 2 Anovas pretendían calibrar la evolución en las calificaciones de los sujetos de control antes y después de la intervención (Media anterior: 23'32, Media posterior: 23'06). Se encontró una F de 0'19, no hallando diferencia significativa. Lo mismo se planteó con los sujetos experimentales en los mismos tiempos (Media anterior: 23'33, Media posterior: 24'09), no hallándose tampoco diferencia significativa, con una F de Anova de 2'96.

Dos últimos Anovas iban a confrontar las calificaciones de los sujetos experimentales y las de los de control antes de la intervención (habíamos previsto en la hipótesis que no se daría diferencia significativa de medias entre las calificaciones de los sujetos experimentales y de los de control) y las de los mismos grupos después de la intervención (suponíamos que habría diferencia significativa a favor de los experimentales).

Antes de la intervención no se dio diferencia significativa de medias: media de los experimentales: 23'33; media de los de control: 23'32. F de Anova: 0'00077.

Después de la intervención tampoco se dio diferencia significativa: media de los experimentales en calificaciones de Lengua y Matemáticas: 24'09. media de los de control: 23'06. F de Anova: 3'14.

A pesar de ello, queremos llamar la atención sobre el hecho de que, partiendo de una situación en que las calificaciones (sus medias) eran casi idénticas en el momento previo a la intervención, después de la misma se ha producido una separación a favor de los sujetos experimentales en más de 1 punto. Sí hay, pues, diferencia entre estos grupos después de la intervención, aunque ésta no sea significativa. Los experimentales han mejorado en sus calificaciones con respecto a los de control con una F de Anova de 3'14, cercana al nivel de significación del 0'05.

No hay que olvidar que, partiendo de datos previos y de nuestros mismos datos de esta investigación, la influencia de la R-I sobre el Rendimiento Académico, aunque existe, es muy moderada. Por eso los resultados de los Análisis de Varianza son bastante congruentes con esta línea general de argumentación. Quizá habíamos sido excesivamente optimistas en nuestras previsiones o quizá habría que confirmarlas con una muestra mayor.

5.5. Estabilidad de las puntuaciones básicas de R-I y validez de constructo de la R-I operacionalizada por el MFF20. Algunas precisiones sobre las puntuaciones «I» y «E».

...*Correlaciones de Latencia-Errores*: se efectuaron para establecer la posible relación existente entre latencia o tiempo previo a la emisión de la respuesta y errores cometidos en la prueba. También se corrobora así la validez de constructo de la R-I, ya que un algo coeficiente «r» de correlación se considera probatorio de la misma.

Tomando datos del Pretest, previo a la intervención para evitar efectos indeseados, encontramos coeficientes de correlación que oscilaban, en cada uno de los 12 grupos, de -0'20 a -0'82. En el macrogrupo formado por todos los sujetos de control el valor fue de -0'69; en el de los experimentales de -0'62, y en todos los sujetos de la muestra, de -0'65.

De acuerdo con la escala de Guilford de valoración de coeficientes de correlación este resultado se considera moderado, pero en el contexto de las investigaciones sobre R-I es un valor elevado, ya que la correlación media oscila, de unas investigaciones a otras, de -0'40 a -0'65, con un promedio de -0'48, según Messer (Messer, 1976). Por tanto, nuestros resultados se sitúan en el punto más elevado. El valor hallado para toda la muestra constata una asociación importante entre tiempo de latencia y errores y sirve para apoyar la validez de constructo del MFF20 como medida de R-I y operacionalizador de la misma a los 12, 13 y 14 años.

Recordemos que investigaciones precedentes realizadas en la misma línea daban valores muy bajos para estas edades, disminuyendo progresivamente desde los 9-10 años (edades consideradas generalmente como las más adecuadas para el marco de las investigaciones sobre R-I), argumentando que, a partir de ellas el constructo iba perdiendo validez. En nuestro caso no ocurre eso, ya que los valores encontrados coinciden con los más altos hallados en investigaciones precedentes para los 9-10 años.

...*Correlaciones de Latencia-Latencia*: van encaminadas a valorar la estabilidad de las puntuaciones de latencia de un pase del test a otro, de cara a validar la consistencia y fiabilidad del constructo.

— Datos del Pretest-Datos del 1.º Postest: al confrontar estos datos vamos a referir aquí únicamente los resultados obtenidos partiendo de los datos de los sujetos de control, a los que no se aplicó el programa de intervención, ya que en los experimentales se esperaba un incremento de tiempo de latencia que, lógicamente, daría unos coeficientes de correlación bajos, si se produjera.

En los 6 grupos de control los coeficientes oscilan de 0'59 a 0'81. Para el macrogrupo formado por todos los sujetos de control se ha encontrado un coeficiente de correlación de 0'63. Estos valores indican una estabilidad aceptable para el intervalo temporal que media entre

ambos pases del test (de tres meses y medio a cuatro meses). Yando encontró un valor de 0'70 para un plazo de 10 semanas (Yando, 1986); Messer encontró una estabilidad de 0'48 y 0'52, respectivamente, para chicos y chicas de otoño a primavera (Messer, 1976). A largo plazo Kagan, en 1965, encontró correlaciones de 0'48 para chicos y chicas, respectivamente, en mediciones efectuadas de 1.º y 2.º grado. Messer, en 1968, halló una estabilidad de latencias en un período de dos años y medio de 0'31 (Messer, 1976).

En el contexto de las investigaciones efectuadas con intervalos temporales similares, el coeficiente hallado para todos los sujetos de control es, pues, elevado.

— Datos del 1.º Posttest-Datos del 2.º Posttest:

En los grupos de control se encuentran de nuevo coeficientes elevados que oscilan de 0'40 a 0'89 según los grupos, dándose para todos los sujetos de control un coeficiente de 0'77. El intervalo temporal ha sido similar al anterior (de tres meses y medio a cuatro meses). Estos coeficientes, muy altos en el contexto de las investigaciones sobre R-I, prueban la estabilidad del constructo para períodos de tiempo de tipo medio y sirven de argumento en favor de su fiabilidad.

Con los grupos experimentales, los coeficientes oscilan de 0'65 a 0'98 grupo a grupo y, para el macrogrupo formado por todos los experimentales se ha dado un coeficiente de 0'82, valores, todos ellos, muy elevados. Esta estabilidad en los experimentales aboga por la eficacia del tratamiento, que ha producido un mantenimiento de elevadas latencias, con un grado de estabilidad todavía mayor que en los sujetos de control.

Para el total de la muestra el valor del coeficiente ha sido de 0'83, de nuevo mucho más elevada que la media de coeficientes hallados en investigaciones previas.

— Datos del Pretest-Datos del 2.º Posttest:

Nos vamos a fijar en los resultados de los sujetos de control, sobre los que no medió intervención, que son los que pueden dar la medida de la estabilidad y consistencia de las latencias, ya que en los experimentales se buscaba incrementarlas haciéndolos más reflexivos.

Grupo a grupo se obtienen coeficientes que oscilan de 0'21 a 0'87. Para el total de los sujetos de control se obtiene un coeficiente de 0'53. Teniendo en cuenta que entre ambos pases del test (1.º y 3.º) median casi ocho meses, hay que hablar de una estabilidad y consistencia aceptables, por encima de la media hallada por otros investigadores para períodos similares.

...*Correlaciones de Errores-Errores*: van encaminadas a valorar la estabilidad de las puntuaciones de error de un pase del test a otro, lo que permitirá validar la consistencia y fiabilidad del constructo R-I operacionalizado por MFF20.

— Datos del Pretest-Datos del 1.º Posttest:

Los coeficientes hallados para los sujetos de control, que son los que servirán de pauta al no haber mediado intervención que, lógicamente, influiría sobre los resultados, oscilan, en los seis grupos, de 0'29 a 0'80, dándose para el macrogrupo de todos los sujetos de control un coeficiente de 0'50. Los datos disponibles al respecto no son coincidentes de unas investigaciones a otras: Messer halló un coeficiente de 0'25 de otoño a primavera (Messer, 1976). Siegelman de 0'43, sin especificar el tiempo transcurrido (Siegelman, 1969). Block et al. de 0'62 y 0'54 (Block et al., 1974). Hay autores que dan una media de 0'30. Ault, sin embargo, sitúa la media en 0'50 (Ault et al., 1976). Existe la constatación general de más elevadas correlaciones para latencia en períodos inferiores a dos años y, por contra, menores para errores y, a partir de dos años, la relación se invierte: los errores son más consistentes que las latencias. De acuerdo con los valores medios de las correlaciones referidas por los autores, el coeficiente hallado en nuestro estudio se sitúa en torno a una media alta y refiere una consistencia aceptable para errores.

— Datos del 1.º Posttest-Datos del 2.º Posttest:

En los seis grupos de control los coeficientes oscilan de 0'43 a 0'92, con un valor, para todos los sujetos de control, de 0'75. Son valores muy elevados en el contexto de las investigaciones sobre R-I y apuntan por una estabilidad y consistencia elevadas para un intervalo temporal de alrededor de 3 meses y medio.

— Datos del Pretest-Datos del 2.º Posttest:

El intervalo temporal es de alrededor de 8 meses. En los seis grupos de control los coeficientes oscilan de 0'37 a 0'67. El macrogrupo de control da un coeficiente de 0'51. De nuevo es un valor elevado para un intervalo temporal tan amplio corroborando una importante estabilidad para las puntuaciones de error.

El MFF20 da lugar a una estabilidad de puntuaciones de error y latencia que se sitúa en los valores más altos en el contexto de las investigaciones sobre R-I (la mayoría de las cuales ha utilizado el MFFT de Kagan). Igualmente son elevados los coeficientes de correlación de latencias con errores. Todo ello apunta por la superioridad del MFF20 sobre el MFFT como operacionalizador de la R-I y prueba la validez de constructo del mismo.

...Puntuaciones de Impulsividad «I» y de Eficiencia «E»:

Salkind y Wright (Salkind y Wright, 1977), saliendo al paso de acertadas críticas sobre los problemas que conllevaba la dicotomización y clasificación de los sujetos por la media de grupo de las puntuaciones de error y de latencia, propusieron asignar a cada sujeto puntuaciones continuas, tanto de estilo (puntuaciones «I» de Impulsividad) como de eficiencia (puntuaciones «E» de Eficiencia), lo que, a su vez, facilitaría la estandarización del test a gran escala. Las puntuaciones de «I» y «E» proceden de las latencias brutas y de las puntuaciones de error del sujeto, transformadas en puntuaciones «z» a través de las siguientes fórmulas:

$$I_i = z_{ei} - z_{li}$$

$$E_i = z_{ei} + z_{li}$$

Las puntuaciones positivas elevadas de «I» son indicativas de impulsividad y las puntuaciones negativas altas lo son de reflexividad.

Las puntuaciones positivas elevadas de «E» indican ineficiencia, mientras que las puntuaciones negativas elevadas de «E» indican eficiencia.

Gjerde, Block y Block reconocieron a estas puntuaciones (Gjerde, Block y Block, 1985) el valor de proporcionar con ellas a cada sujeto una puntuación continua, pero critican que combinan desafortunadamente a los sujetos Rápidos-Exactos y Lentos-Inexactos asignando a los dos grupos puntuaciones equivalentes intermedias.

Eso no es cierto en absoluto: si bien los dígitos que obtienen el «I» y «E» estos dos grupos pueden ser similares (cosa que, de todas formas, habría que demostrar), los Lentos-Inexactos tenderán a obtener cifras positivas, indicativas de ineficiencia, y los Rápidos-Exactos negativas, indicativas de eficiencia. Lo mismo ocurrirá con las puntuaciones «I» de impulsividad-reflexividad. Los gráficos que, en nuestra tesis doctoral, realizamos al efecto, ilustran estas afirmaciones con claridad (Gargallo, 1989).

Para no extendernos demasiado vamos a corroborar estas afirmaciones con dos individuos de uno de los grupos: el uno, el número 9, había sido clasificado como Lento-Inexacto, el número 3, como Rápido-Exacto.

Número 9: Lento-Inexacto: puntuación total de errores obtenidos en el 2.º pase del test: 7; media de latencia empleada por el mismo al resolver la prueba: 30'4 segundos.

$$z = \frac{x - \bar{x}}{\sigma} \quad z_{ei} = 0'42186 \quad z_{li} = 0'00440$$

$$I_i = z_{ei} - z_{li} = 0'41746$$

$$E_i = z_{ei} + z_{li} = 0'42626$$

Número 3: Rápido-Exacto: puntuación total de errores obtenidos en el 2.º pase del test: 1; media de latencia: 28'3 segundos.

$$z_{ei} = -0'19111 \quad z_{li} = 0'00008$$

$$I_i = z_{ei} - z_{li} = -0'19203$$

$$E_i = z_{ei} + z_{li} = -0'19119$$

Los resultados son nítidos: el Lento-Inexacto obtiene puntuaciones positivas, indicativas de ineficiencia e impulsividad y el Rápido-Exacto las obtiene negativas, indicativas de eficiencia y reflexividad.

Queremos precisar que estas puntuaciones «I» y «E» partían de las medias y las sigmas de toda la muestra, con lo que resultaban notablemente más suaves que si hubieran partido de la propia media del grupo, en cuyo caso las diferencias hubieran sido mayores.

Además se puede observar que la media de latencia difiere poco en ambos sujetos; el Rápido-Exacto tiene 28'3 segundos y el Lento-Inexacto 30'4 segundos. Si la puntuación de latencia del Rápido-Exacto fuera menor, ello implicaría una puntuación de «I» ligeramente menor y una puntuación «E» algo mayor que la obtenida, pero en ambos casos negativas: sería ligeramente menos reflexivo y algo más eficiente.

Si el Lento-Inexacto tuviera una puntuación de latencia mayor y se alejase más de la media del grupo por exceso, la puntuación positiva de «E» todavía sería mayor, señalando mayor ineficiencia, y la de «I» menor, indicando algo más de reflexividad que en el caso actual.

Con ello la crítica de Gjerde et al. antes aludida queda completamente refutada.

6. Conclusiones y recomendaciones

Tal y como hemos venido reseñando las previsiones de la primera parte de la hipótesis se han cumplido en grado altamente satisfactorio.

El programa ha sido sumamente eficaz en el logro de los objetivos propuestos de cara a incrementar la reflexividad de los sujetos y su eficacia para resolver la tarea y en mantenimiento de estos resultados positivos en el 2.º Postest (perdurabilidad y estabilidad).

El importante N del total de la muestra (201 sujetos) y el hecho de haber experimentado el programa en 5 colegios diferentes con distintos experimentadores, junto con el éxito del programa y la perdurabilidad y estabilidad de resultados, permite apostar por su generalización y por la validez del programa cuando se aplique a otros grupos de nivel similar de edad, conocimientos y desarrollo cognitivo.

No se encontraron diferencias sexuales significativas en la ejecución del MFF20: varones y hembras tenían niveles similares de reflexividad-impulsividad.

Encontramos diferencias significativas en la ejecución del test en las tres edades del estudio: 12, 13 y 14 años, con un incremento de reflexividad de 12 a 13 años y con un incremento de impulsividad de 13 a 14, datos que corroboran hallazgos previos nuestros y que contradicen otras afirmaciones clásicas en autores que han trabajado con el MFFT. Estos datos contradicen, también, la tesis del efecto de techo para la R-I a los 9-10 años o a los 10-12 y sirven para constatar que la R-I sigue creciendo, al menos hasta los 13 años.

Constatamos, también, una mayor reflexividad en los sujetos de nuestra muestra que en las de Salkind y Cairns y Cammock. Sucesivos estudios deberían confirmar si, según parece desprenderse, los niños españoles son más reflexivos que los estadounidenses y norirlandeses.

Encontramos una moderada influencia de la R-I sobre el Rendimiento Académico, una primera diferencia significativa entre las calificaciones de reflexivos e impulsivos a favor de los primeros y, luego, una mejora en las calificaciones de los sujetos sometidos a tratamiento experimental, que no llegó a ser significativa. Tampoco los sujetos que incrementaron su reflexividad en el transcurso del tratamiento mejoraron significativamente sus calificaciones, ni los que se hicieron más eficientes. Tal vez se debió al bajo N de los grupos enfrentados. Queremos insistir en que los sujetos experimentales y los de control tenían medias de calificaciones casi idénticas en el momento previo a la intervención y que, después de la misma, los experimentales superaban en más de 1 punto a la media de calificaciones de los de control. No obstante el Análisis de Varianza no dio diferencia significativa de media.

Hallamos unos coeficientes de correlación, tanto al correlacionar latencias con errores como al hacerlo con latencias-latencias y errores-errores, con valores que se situaban en los límites superiores de los hallados para el MFFT o que los superaban, abonando nuestra hipótesis de superioridad del MFF20 como operacionalizador de la R-I por la mayor consistencia y fiabilidad de resultados. también refutamos las afirmaciones de Gjerde, Block y Block relativas las puntuaciones «I» y «E», de Impulsividad y Eficiencia, ya que no se da la confusión de Rápidos-Exactos y Lentos-Inexactos, que ellos sugieren, en estas puntuaciones.

Encontramos también que era el componente intervención el que producía mayor diferenciación entre los grupos, frente a otras variables también analizadas en este artículo.

Todo ello ha servido, pues, para corroborar algunas de las afirmaciones clásicamente mantenidas sobre diversas variables importantes en el contexto de la R-I y para cuestionar otras comúnmente sostenidas.

Queremos añadir que fue la observación de la existencia de un abundante número de alumnos impulsivos en el aula escolar la que nos impulsó a sumergirnos en esta investigación. Los impulsivos cometen más errores que los reflexivos y tienen más problemas en su andadura escolar y menor éxito académico. Para que esto no desemboque en una actitud negativa frente a la actividad escolar por un sentimiento básico de impotencia ante el fracaso acumulado, hay que modificar su impulsividad y reconducirla hacia una mayor reflexividad. Esto se puede lograr, como se ha podido comprobar, en un marco intervencionista, con programas de acción educativa similares a éste, que ya ha demostrado su utilidad y eficacia.

Es cierto que los estilos cognitivos son consistencias con alto grado de estabilidad y, por tanto, difíciles de cambiar, aunque nosotros creemos que esto es posible. De todas formas es evidente, como lo demuestra este trabajo, que se puede dotar a los niños de instrumentos y estrategias cognitivos adecuados de tal forma que se dé una mayor adaptación entre sus estrategias de procesamiento de la información habituales y las demandas que la tarea académica presenta que, en ocasiones, no coinciden con aquellas, con un incremento de tiempo de latencia que se va a emplear en escudriñarla más cuidadosamente.

Un aspecto a reseñar del programa es su viabilidad y funcionalidad, sin especiales requerimientos de formación para su aplicación. Es sencillo, se inserta perfectamente en la marcha normal de la clase sin truncarla, precisa de poco tiempo, es motivador para los niños y genera

buenos resultados. Todo ello puede favorecer, desde nuestro punto de vista, el uso del mismo por parte del profesorado, que constata como nosotros mismos las dificultades de los niños impulsivos en la marcha normal del aula y no tiene herramientas adecuadas para corregir estos problemas. Nosotros ya podemos poner una a su disposición.

Dirección del autor: Bernardo Gargallo López, c/ San José de la Montaña, n.º 26, pta. 9, 46008 Valencia.

Fecha de recepción de la versión definitiva de este artículo: 25.II.1991.

BIBLIOGRAFÍA

- AULT, R. L.; MITCHELL, C. y HARTMAN, D. P. (1976) Some methodological problems in reflection-impulsivity research, *Child Development*, 47, pp. 227-231.
- BENTLER, P. M. y McCLAIN, J. (1976) A multitrait multimethod analysis of reflection-impulsivity, *Child Development*, 47, pp. 218-226.
- BLOCK, J. (1987) Misgivings about the Matching Familiar Figures Test: Premature or Overdue?, *Developmental Psychology*, 23:1, pp. 740-741.
- BLOCK, J.; BLOCK, J. H. y HARRINGTON, D. M. (1974) Some misgivings about the Matching Familiar Figures Test as a measure of reflection-impulsivity, *Developmental Psychology*, 10, pp. 611-632.
- (1975) Comment on the Kagan-Messer replay, *Developmental Psychology*, 11, pp. 249-252.
- BLOCK, J.; GJERDE, P. F. y BLOCK, J. H. (1986) More misgivings about the Matching Familiar Figures Test as a measure of Reflection-Impulsivity: Absence of Construct Validity in Preadolescence, *Developmental Psychology*, 22: 6, pp. 280-831.
- BRIGGS, C. H. (1968) *An experimental study of reflection-impulsivity children* (Doctoral dissertation, University of Minnesota, 1966) (Dissertation Abstracts, 1968, 28, 3891-B, University Microfilms, n. 68-1610).
- BRIGGS, C. H. y WEINBERG, R. A. (1973) Effects of reinforcement in training children's conceptual tempo, *Journal of Educational Psychology*, 65, pp. 383-394.
- CAIRNS, E. y CAMMOCK, T. (1978) The development of a more reliable version of the Matching Familiar Figures Test, *Developmental Psychology*, 5, pp. 555-560.
- (1984) The development of reflection-impulsivity: further data, *Personality Individual Differences*, 5:1, pp. 113-115.
- (1989) *The 20-Item Matching Familiar Figures Test: Technical data*. Manuscrito no publicado, enviado por el prof. Cairns al autor de este artículo en julio de 1989.
- CASTILLEJO, J. L. y GARGALLO, B. (1989) Un programa de intervención para mejorar la reflexividad en preadolescentes (8.º EGB Primaria), *Revista Española de Pedagogía*, 184, pp. 539-555.
- COHEN, S. y PRZYBYCIEN, C. A. (1974) Some effects of sociometrically selected peer models on the cognitive styles of impulsive children, *Journal of Genetic Psychology*, 124, pp. 213-220.
- DEBUS, R. L. (1970) Effects of brief observation of model behavior on conceptual tempo of impulsive children, *Developmental Psychology*, 2, pp. 22-32.

- (1976) *Observational Learning of reflective strategies by impulsive children*. Paper prepared for the Symposium on Observational learning. XXI Congrès Internationale de Psychologie. Julio, 1976.
- DENNEY, D. R. (1972) Modeling effects upon conceptual style and cognitive tempo, *Child Development*, 43, pp. 105-119.
- EGELAND, B. (1974) Training impulsive children in the use of more efficient scanning strategies, *Child Development*, 45, pp. 165-171.
- EGELAND, B. y WEINBERG, R. A. (1976) The Matching Familiar Figures Test: a look to its psychometric credibility, *Child Development*, 47, pp. 483-491.
- ERRICKSON, E. A.; WYNE, M. C. y ROUTH, D. K. (1973) A response-cost procedure for reduction of impulsive behavior of academically handicapped children, *Journal of Abnormal Child Psychology*, 1, pp. 350-357.
- GAINES, P. D. (1971) *The modification of attentional strategies in children* (Report, N.º 1, Developmental Program, Department of Psychology) Unpublished manuscripts, March, 1971 (Available from. L. Nadelman, Department of Psychology, University of Michigan, Ann Arbor, Michigan 48104).
- GARGALLO, B. (1985) *El estilo cognitivo «Reflexividad-Impulsividad» y su modificabilidad. Un programa de intervención para 5.º y 6.º de EGB*. Tesina de licenciatura no publicada, Universidad de Valencia.
- (1987) La Reflexividad como objetivo educativo: un programa de acción educativa, en VARIOS *Investigación educativa y práctica escolar. Programas de acción en el aula* (Madrid, Santillana/Aula XXI).
- (1989) *El estilo cognitivo «Reflexividad-Impulsividad». Su modificabilidad en la práctica educativa. Un programa de intervención para 8.º de EGB* (tesis doctoral publicada en microficha, Universidad de Valencia).
- GJERDE, P. F.; BLOCK, J. y BLOCK, J. H. (1985) Longitudinal consistency of Matching Familiar Figures Test Performance from Early Childhood to Preadolescence, *Developmental Psychology*, 21:2, pp. 262-271.
- JELSMA, O. y PETERS, J. M. (1989) Practice Schedule and Cognitive Style Interaction in Learning a Maze Task, *Applied Cognitive Psychology*, 3:1, pp. 73-83.
- KAGAN, J. (1965a) *Matching Familiar Figures Test* (Cambridge, Massachusetts, Harvard University).
- (1965b) Reflection-Impulsivity and reading ability in primary grade children, *Child Development*, 36, pp. 609-628.
- (1965c) Impulsive and reflective children: significance of conceptual tempo, en KRUMBOLTZ, J. A. (Ed.) *Learning and the educational process* (Chicago, Rand McNally).
- (1966) Reflection-Impulsivity: The generality and dynamics of conceptual tempo, *Journal of Abnormal Psychology*, 71, pp. 17-24.
- (1987) Misgivings about the Matching Familiar Figures Test: A brief replay to Block, Gjerde and Block (1986), *Developmental Psychology*, 23:5, pp. 738-739.
- KAGAN, J.; PEARSON, L. y WELCH, L. (1966) Modifiability of an impulsive conceptual tempo, *Journal of Educational Psychology*, 57, pp. 359-365.
- KREITLER, H. y KREITLER, S. (1976) *Cognitive Orientation and Behavior* (New York, Springer).
- (1982) Cognitive orientation: expanding the scope of behavior prediction, en MAHER, B. y MAHER, W. B. (Eds.) *Advances in Experimental Personality Research* (vol. 11) (New York, Academic Press).
- LEWIS, M. (1971) Sex differences in cognitive styles: A rejoinder, *Perceptual and Motor Skills*, 33, pp. 1.006.

- LOPER, A. B.; HALLAND, D. y MCKINNEY, J. D. (1982) The effect of reinforcement for global or analytic strategies on the performance of reflective and impulsive children, *Journal of Experimental Child Psychology*, 33, pp. 55-62.
- LÖSEL, F. (1980) On the differentiation of cognitive reflection-impulsivity, *Perceptual and Motor Skills*, 50, pp. 1.311-1.324.
- MEICHENBAUM, D. H. (1971) *The nature and modification of impulsive children: training impulsive children to talk themselves*. Manuscrito inglés (Adaptación de un trabajo presentado para la conferencia de 1971 de la SRCE celebrado en Minneapolis, Minnesota).
- (1977) *Cognitive-behavior modification: An integrated approach* (New York, Plenum Press).
- (1981) Una perspectiva cognitivo-comportamental del proceso de socialización, *Análisis y Modificación de Conducta*, 7:14 y 15, pp. 85-109.
- MEICHENBAUM, D. H. y ASARNOW, J. (1979) Cognitive behavior modification and metacognitive development: Implications for the classroom, en KENDALL, P. y HOLLON, J. (eds.) *Cognitive behavioral interventions: Theory research and procedures* (New York, Academic Press).
- MEICHENBAUM, D. M. y GOODMAN, J. (1969) Reflection-impulsivity and verbal control of motor behavior, *Child Development*, 40, pp. 785-797.
- (1971) Training impulsive children to talk to themselves: A means of developing self-control, *Journal of Abnormal Psychology*, 77, pp. 115-126.
- MESSER, S. B. (1976) Reflection-impulsivity: a review, *Psychological Bulletin*, 83:6, pp. 1.026-1.052.
- NAGIE, R. J. y THWAITE, B. C. (1979) Are learning disabled children more impulsive? A comparison of learning disabled and normal-achieving children on Kagan's Matching Familiar Figures Test, *Psychology in the Schools*, 16, pp. 351-355.
- RIDBERG, E. H.; PARKE, R. D. y HETHERINGTON, E. M. (1971) Modification of impulsive and reflective cognitive styles through observation of film mediated models, *Developmental Psychology*, 5, pp. 369-377.
- SALKIND, N. J. (1978) The development of norms for the Matching Familiar Figures Test, *JSAS, Catalog of Selected Documents in Psychology*, 8, 61.
- SALKIND, N. J. y NELSON, C. F. (1980) A note on the developmental nature of reflection-impulsivity, *Developmental Psychology*, 16:3, pp. 237-238.
- SALKIND, N. J. y WRIGHT, J. C. (1977) The development of reflection-impulsivity and cognitive efficiency: an integrated model, *Human Development*, 20, pp. 377-387.
- SCHACHAR, R. y LOGAN, G. D. (1990) Impulsivity and Inhibitory Control in Normal Development and Childhood Psychopathology, *Developmental Psychology*, 26:5, pp. 710-720.
- SCHWEBEL, A. (1966) Effects of impulsivity on performance of verbal tasks in middle and lower-class children, *American Journal of Orthopsychiatry*, 36, pp. 12-21.
- SIEGELMAN, E. (1969) Reflective and impulsive observing behavior, *Child Development*, 40, pp. 1.213-1.222.
- WALCZYK, J. J. y HALL, V. C. (1989) Is the Failure to Monitor Comprehension an Instance of Cognitive Impulsivity?, *Journal of Educational Psychology*, 81:3, pp. 294-298.
- WARD, W. (1968) Reflection-impulsivity in kindergarten children, *Child Development*, 39, pp. 867-874.
- WEINBERG, R. A. (1969) *The effects of different types of reinforcement in training a reflective conceptual tempo* (Doctoral Dissertation, University of Minnesota, 1969) (Dissertation Abstracts, 1969, 29, 2578-A, University Microfilms n.º 69-1560).

- YANDO, R. y KAGAN, J. (1968) The effect of teacher tempo on the child, *Child Development*, 39, pp. 27-34.
- ZAKAY, D.; BAR-EL, Z. y KREITLER, S. (1984) Cognitive Orientation and changing the impulsivity of children, *British Journal of Educational Psychology*, 54, pp. 40-50.
- ZUCKER, J. y STRIKER, G. (1968) Impulsivity-reflectivity in preschool head-start and middle class children, *Journal of Learning Disabilities*, 1, pp. 24-30.

SUMMARY: REFLECTION-IMPULSIVITY. SOME IMPORTANT VARIABLES. LAST FINDINGS. PERDURABILITY OF RESULTS OBTAINED THROUGH AN INTERVENTION PROGRAMME.

We found a substantial increasing of reflection in 6 experimental groups opposite to 6 control groups in a 1st and in a 2nd Posttests. Both groups showed similar levels of R-I in a Pretest. The results had perdurability and stability for a long time after an intervention programme. There was no significative difference between boys and girls. There was significative difference of means between 12, 13 and 14 years old. Reflection increased at 13 but not at 14. MFF20 is adequate to these ages. Spanish children were more reflective than American and Northirish children (Salkind's and Cairns and Cammock's data). There was a moderate influence of R-I concerning to academic efficiency. Correlations between latency and errors gave high «r» coefficients confirming the superiority of MFF20 in front of the MFF test of Kagan. «I» and «E» scores are adequate and there is no confusion between slow-inaccurate and fast-accurate subjects as Gjerde et al. say when you use them.

KEY WORDS: Reflection-Impulsivity/Reflectivity-Impulsivity. Reflection increasing. Reflection's modifiability, stability and perdurability. Sex and age performance of MFF20. «I» and «E» scores of R-I. Reflection-impulsivity and academic efficiency.