

Problemas de impulsividad e inatención en el niño



MINISTERIO
DE EDUCACIÓN,
CULTURA Y DEPORTE

Centro de
Investigación y
Documentación Educativa

c i d e

PROBLEMAS DE IMPULSIVIDAD E INATENCIÓN EN EL NIÑO

Propuestas para su evaluación

**Mateu Servera Barceló
Maria Rosario Galván Pascual**

PROBLEMAS DE IMPULSIVIDAD E INATENCIÓN EN EL NIÑO



Número: 152

Colección: INVESTIGACIÓN



MINISTERIO DE EDUCACIÓN, CULTURA Y DEPORTE
SECRETARÍA GENERAL DE EDUCACIÓN Y FORMACIÓN PROFESIONAL
Centro de Investigación y Documentación educativa (C.I.D.E.)

Edita:

© SECRETARÍA GENERAL TÉCNICA

Subdirección General de Información y Publicaciones

N.I.P.O.: 176-01-155-9

I.S.B.N.: 84-369-3502-0

Depósito legal: M-43740-2001

Imprime: Gráficas Don Bosco S.L.

INDICE

PARTE TEÓRICA: EL ESTILO IMPULSIVO Y LA CAPACIDAD DE ATENCIÓN SOSTENIDA

1. Introducción: el proyecto IMAT.....	9
2. La naturaleza multicomponente de la dimensión "impulsividad"	11
2.1. La naturaleza motora, social y cognitiva	11
2.2. El estilo cognitivo reflexividad-impulsividad	14
2.2.1. Los estilos cognitivos y los estilos de aprendizaje.....	14
2.2.2. La dimensión reflexividad-impulsividad	18
2.3. La evaluación del estilo impulsivo: los tests de emparejamiento de figuras familiares (MFF).....	30
2.3.1. Aspectos psicométricos de los MFF.....	35
2.3.2. Sistemas de clasificación en los MFF: el modelo integrado de Salkind y Wright	46
2.4. El aprendizaje y el comportamiento del "niño impulsivo"	54
3. La naturaleza multicomponente de la atención.....	61
3.1. Conductas de inatención y déficit de atención en el niño	61
3.2. El funcionamiento atencional.....	68
3.2.1. Definición y concepto.....	68
3.2.2. Los tipos de atención: selectiva, dividida y sostenida	73
3.3. La evaluación de la atención sostenida: las tareas de ejecución continua y las tareas de vigilancia	78
3.3.1. Características generales y tipos de medida en las tareas de vigilancia y ejecución continua	81
3.3.2. Sistemas de clasificación: la teoría de detección de señales	87
3.4. El trastorno por déficit de atención con y sin hiperactividad.....	94

PARTE EXPERIMENTAL

4. Los objetivos del trabajo	99
5. Metodología	102
5.1. Sujetos	102
5.2. Las pruebas de evaluación	104
5.2.1. La Tarea de Emparejamiento de Figuras (TEF-IMAT)	104
5.2.2. La Tarea de Vigilancia (TVI-IMAT)	105
5.2.3. La Escala de Comportamiento y Rendimiento en el Aula (ECRA)	108
5.3. Procedimiento	112
6. Análisis de resultados	118
6.1. Resultados descriptivos y comparación por cursos y por sexo	118
6.1.1. Datos de la TEF-IMAT	118
6.1.2. Datos de la TVI-IMAT	119
6.2. Análisis de fiabilidad y consistencia interna	123
6.2.1. Datos de la TEF-IMAT	123
6.2.2. Datos de la TVI-IMAT	124
6.3. La relación entre el estilo impulsivo y la falta de atención sostenida ...	124
6.4. Las implicaciones de la impulsividad cognitiva en el aula	129
6.5. Las implicaciones de la atención sostenida en el aula	134
7. Conclusiones y perspectiva de futuro	139
7.1. Principales conclusiones en relación a los objetivos del trabajo	139
7.2. El futuro de la evaluación objetiva de la impulsividad y la inatención en el aula	151
Referencias bibliográficas	159

PARTE TEÓRICA: EL ESTILO IMPULSIVO Y LA CAPACIDAD DE ATENCIÓN SOSTENIDA

1. INTRODUCCIÓN: EL PROYECTO IMAT

En el año 1997 los autores de la presente monografía pusieron en marcha el proyecto IMAT, con el objetivo de profundizar en el estudio de los problemas de impulsividad e inatención en el niño. Mateu Servera, junto con el profesor Xavier Bornas, había abordado el tema de la impulsividad en distintos proyectos de investigación a lo largo de toda la década. Precisamente uno de estos primeros proyectos fue subvencionado por el CIDE en 1991 centrado en la modificación del estilo cognitivo impulsivo. La experiencia, los datos acumulados y especialmente el contacto muy intenso con el ámbito escolar de nuestra comunidad nos han seguido impulsando en este campo, ampliándolo también al ámbito clínico, concretamente a través de la Unidad de Hiperactividad del Departamento de Psicología de la UIB. Todo ello nos llevó al convencimiento de que los problemas de impulsividad e inatención eran cruciales en la enseñanza primaria y constituían uno de las principales preocupaciones de educadores, padres y profesionales.

Tal vez por la presión que ejerce esta preocupación ampliamente compartida se detecta, desde hace muchos años ya, un gran afán (del que nosotros en un primer momento también fuimos partícipes) por buscar soluciones rápidas y eficaces: desde las farmacológicas hasta toda la gama psicoeducativa imaginable. Pero el hecho de disponer de bastantes herramientas para combatir la impulsividad y la inatención infantil no ha impedido que éstos sigan siendo problemas de primer orden en el aula. Desde luego las razones para ello son múl-

tiples, y tal vez podríamos empezar por enmarcarlas en la tortuosa relación entre la investigación y la práctica educativa, pero también es posible centrarlas en aspectos más concretos. Creemos que uno de los problemas de las técnicas de tratamiento de la impulsividad y la inatención es que se han desarrollado sin atender al hecho de que no sólo hay un tipo de impulsividad y un tipo de inatención. La psicología experimental tendría mucho que decir al respecto, pero la psicología aplicada, tanto la clínica como la educativa, se ha mantenido a cierta distancia de los problemas de conceptualización. El resultado, a nuestro juicio, es un panorama confuso donde expertos y profanos juegan una y otra vez con los términos «*impulsivo*», «*déficit de atención*» e «*hiperactivo*» refiriéndose a conductas, síndromes o alteraciones muy diversas.

Así nació el proyecto IMAT, con implicación de profesores universitarios y maestros de educación primaria. Su objetivo fundamental es, obviamente, ayudar a la solución de problemas de inatención e impulsividad, pero para ello se decidió partir de un planteamiento científico, en el sentido de definir claramente cuál era el problema con el que nos enfrentábamos. El primer paso era analizar conceptualmente las dimensiones impulsividad e inatención y después proponer sistemas de evaluación objetivos, complementarios de otros más habitualmente utilizados. Y eso es precisamente lo que hemos intentado en esta monografía.

En última instancia el lector tendrá a su disposición dos pruebas de evaluación, una del estilo cognitivo impulsivo y otro de la capacidad de atención sostenida, prácticamente listas para su uso en el ámbito aplicado, tanto para evaluar como para controlar la eficacia de un tratamiento. La primera de estas pruebas ya ha sido publicada, en una versión profesional, por el grupo Albor-Cohs (Servera y Llabrés, 2000) y la segunda esta en proceso, aunque puede utilizarse ya en el ámbito de la investigación. El lector puede consultar estas dos pruebas en internet (<http://www.uib.es/facultat/psicologia/downloads/imat.exe>) o ponerse en contacto con el primer autor, Mateu Servera, para obtener más información al respecto (mateus@uib.es).

En esta monografía encontrará todo el proceso conceptual y metodológico que hemos seguido para llegar a poner a disposición de todos estas pruebas.

2. LA NATURALEZA MULTICOMPONENTE DE LA DIMENSIÓN «IMPULSIVIDAD»

2.1 La impulsividad motora, social y cognitiva

En nuestra revisión de 1996 sobre la impulsividad (Bornas y Servera, 1996) dedicábamos un primer capítulo introductorio relativamente amplio simplemente a clarificar a qué tipo de impulsividad nos referiríamos a lo largo del trabajo. No era una novedad; en su magnífico trabajo de revisión de 1993 Dickman, titulado *«Impulsivity and information processing»*, empezaba de modo similar. Curiosamente muchos creyeron que nuestro libro, subtulado *«un enfoque cognitivo-conductual»*, trataría la impulsividad conductual o social, mientras el de Dickman parece más destinado a la impulsividad cognitiva, y cuando el lector los aborda se da cuenta que prácticamente es al revés. Y es que pocos temas en Psicología parecen tan relevantes y a la vez tan polémicos como éste.

No hay una definición única de impulsividad porque, aún con polémicas, mayoritariamente se reconoce la existencia de distintos tipos de impulsividad, a veces sólo tangencialmente relacionados (el porqué se les sigue denominando con el mismo nombre es uno de estos misterios insondables en nuestra disciplina). En realidad, de todos modos, no es del todo cierto que la clasificación que hemos propuesto nosotros: motora, social y cognitiva, sea la que nos parezca más acertada. Eso son, como para cualquier otra dimensión comportamental, canales de respuesta que nos sirven de entrada para plantear el tema. A nosotros nos parece que hay datos suficientes para hablar de una impulsividad manifiesta (con un componente motor y otro social), una impulsividad personalógica (con influencias de estilo cognitivo, por un lado, y por otro, del

marco biofactorial) y, finalmente, una impulsividad del procesamiento de la información o cognitiva (con una conceptualización dentro de las teorías del pensamiento y dentro de los modelos cognitivo-conductuales).

La impulsividad manifiesta arranca de los trabajos de Skinner sobre el tema del «*control de estímulos*», que Barkley ha recuperado en sus modernas teorías explicativas de la hiperactividad (Barkley, 1990; 1998). Curiosamente éste es también un primer punto de encuentro de la impulsividad y la inatención. Los dos reflejan un problema de control de estímulos porque la atención se entiende como una baja relación entre un estímulo y la conducta esperada del individuo, mientras la impulsividad se entiende, por un lado, como una incapacidad para demorar reforzadores (baja tolerancia a la frustración y/o poca resistencia a la tentación) y, por otro, como una falta de consideración de las condiciones estimulares presentes (precipitación e incapacidad de previsión de consecuencias). Este tipo de impulsividad, conocida hoy en día como un trastorno del bajo control, es la más evidente en trastornos hiperactivos y algunos trastornos de conducta graves. Su naturaleza se basa en modelos que interaccionan aspectos hereditarios, funciones neuropsicológicas y factores de aprendizaje. Sin duda tiene un gran interés clínico, pero no es el objetivo de nuestro trabajo.

La impulsividad personalógica es la que deriva, por un lado, de la teoría de los estilos cognitivos y, por otro lado, de las teorías biofactoriales de la personalidad. En el siguiente apartado abordaremos el estilo cognitivo reflexividad-impulsividad, o mejor dicho abordaremos su reformulación, puesto que son pocos los datos hoy en día que sigan avalando la posibilidad de hablar de un estilo impulsivo en términos ortodoxos. Por su parte, la impulsividad sí continua jugando un papel clave en prácticamente todas las teorías de la personalidad vigentes. Dickman (1993) define esta impulsividad como una tendencia a actuar con una menor previsión de las consecuencias de nuestros actos, y distingue entre una impulsividad funcional y otra disfuncional. La primera se refiere a conductas de falta de previsión que han recibido contingencias positivas asiduamente (en otras palabras, han sido reforzadas por el

medio); es la típica de personas muy activas, buscadores de riesgos y sensaciones, que de algún modo poseen sistemas que les protegen de los posibles errores normalmente asociados a su conducta impulsiva. La disfuncional, en cambio, es la impulsividad que aún habiendo recibido mucha estimulación aversiva se mantiene; es la de las personas que a la precipitación normalmente añaden errores o resultados negativos. En términos factoriales, la impulsividad personológica normalmente es uno de los dos componentes de la dimensión «*extraversión*» (el otro es la sociabilidad), y está presente en los modelos y las pruebas de evaluación (tests, autoinformes, escalas, etc.) de Eysenck, Barrat, Costa y McRae, Gray, etc., es decir, en los principales autores de teorías de la personalidad.

La impulsividad manifiesta y la personológica pueden tener puntos de conexión, tal vez una es continuación de la otra en la edad adulta en algunos casos, pero en otros no es así. Los trastornos del bajo control no tratados se pueden relacionar, en la edad adulta, con delincuencia, agresividad y problemas de adaptación social, mientras la impulsividad personológica, incluso la disfuncional, puede causar problemas puntuales más o menos importantes a la persona que la presenta, pero no tiene porqué llegar al ámbito clínico.

La impulsividad del procesamiento de la información es la que a nosotros nos interesa en este trabajo. Es un tipo de impulsividad ligada al ámbito cognitivo, más específicamente al afrontamiento y resolución de problemas. Sabemos que alcanza su mayor relevancia entre los 6-12 años, puesto que va muy ligada a aspectos de maduración que tal vez hagan que en la adolescencia y en sujetos adultos su repercusión sea menor. No es que no tenga continuidad, pero no está nada claro que esa continuidad sea la personológica, más bien se diluye por el desarrollo de otros mecanismos cognitivos compensatorios que la hacen menos evidente. En los modelos explicativos de la acción de pensar esta impulsividad hace referencia a un sesgo en la actuación del sujeto que altera cualquiera de los procesos básicos de la resolución del problema: su identificación, su definición, la generación de alternativas de solución, la previsión de consecuencias, la toma de decisión, etc. Dentro del

modelo cognitivo-conductual, aunque más aplicado, sigue manteniendo una estructura muy similar, sólo que en este caso el trabajo de investigación se ha orientado más a su tratamiento que a su conceptualización (véase Bornas, 1994; Bornas y Servera, 1996; Servera, 1992).

La impulsividad cognitiva se puede definir como una tendencia del sujeto a precipitar sus respuestas, especialmente en tareas que explícita o implícitamente conllevan incertidumbre de respuesta, y a cometer más errores. Es, pues, una impulsividad contextualizada, en comparación con las otras dos más generales, pero ese contexto no es tan limitado como a primera vista pudiera parecer: la mayoría de tareas y actividades del aprendizaje escolar requieren una aproximación reflexiva para su correcta asimilación y ejecución. La impulsividad cognitiva es, pues, una impulsividad básicamente infantil. En definitiva, la impulsividad cognitiva comparte con las otras dos la idea de precipitación y falta de consideración de las consecuencias, pero se diferencia de la manifiesta en que no tiene porqué ir ligada a problemas de falta de control motor, y de la personológica en que no tiene porqué demostrarse en el espectro de funcionamiento social de la persona. Al menos así lo atestiguan los datos que, en lugar de presentarnos un constructo de «*impulsividad*» amplio y sólido en sus tres canales de respuesta, se empeñan una y otra vez en defender una relativa independencia de las tres dimensiones aquí especificadas.

Nosotros nos centraremos básicamente en el estilo cognitivo (o del procesamiento de la información) reflexividad-impulsividad, si bien con alguna alusión necesaria a los otros tipos.

2.2 El estilo cognitivo reflexividad-impulsividad

2.2.1 Los estilos cognitivos y los estilos de aprendizaje

La dimensión reflexividad-impulsividad se estudia, se desarrolla y se explica normalmente dentro del campo de los estilos cognitivos. Sin embargo, a pesar de que cierta base para esta inclusión, son muchos los aspectos que no encajan. En nuestro caso, personalmente, preferi-

mos referirnos a la reflexividad-impulsividad como un estilo o estrategia de aprendizaje que, en su conceptualización más formal, entroncaría con una de las líneas de estudio de los estilos de procesamiento de la información.

En otro lugar expusimos más ampliamente el concepto de estilo cognitivo y la definición e implicaciones de los dos estilos más conocidos: la dependencia-independencia de campo (DIC) y la reflexividad-impulsividad (R-I) (véase, Servera, 1997). Allí ya discutíamos las dificultades para seguir manteniendo a la R-I dentro de este ámbito y, en todo caso, la decisión final vendría claramente influenciada por el propio concepto de «*estilo cognitivo*» que adoptásemos. Quiroga y Sánchez (1987) destacaban la presencia de tres tipos de definiciones: las que conciben a los estilos cognitivos como variables mediadoras, pero no integradoras, de la actividad cognitiva, las que los conciben como actividad mediadora e integradora de esta actividad, y las que abarcan su influencia a otros aspectos «*no cognitivos*» como el comportamiento, la personalidad, la respuesta emocional, etc. Las autoras concluían que los estilos cognitivos son «*un constructo hipotético superordinal, desarrollado para describir y, en su caso, explicar las diferencias individuales observadas en alguno alguno de los procesos que medían entre el estímulo y la respuesta, integrando los aspectos cognitivos y no cognitivos de la persona*» (Quiroga y Sánchez, 1987, p. 62).

En su tiempo se fueron definiendo bastantes estilos cognitivos, algunos solapados, como, por ejemplo, el de enfoque-exploración, el de serialistas-holistas, el de nivelamiento-agudización, control flexible-rígido, etc. En realidad todos ellos intentaban cubrir las expectativas que había creado la DIC de Herman Witkin desde mediados de los años cincuenta, pero prácticamente ninguno, ni siquiera la R-I lo ha conseguido, si tomamos como referencia la anterior definición. Desde este punto de vista un estilo cognitivo cubre el amplio tramo que va desde las teorías de la personalidad hasta las teorías de la percepción, y debe caracterizarse por lo siguiente: describir modos o formas de manejar la información más que los productos finales de esta acción, estas formas de pensar no están sujetas a juicios de valor (serán más o menos adaptativas

en función de la tarea), suelen distribuirse a lo largo de un continuo bipolar, presentan consistencia interindividual e intrasituacional y tienen referentes en el campo cognitivo-educativo y en el comportamental-social del individuo. Son fruto de la interacción entre factores predisponentes y aprendidos, aunque en principio son difíciles de modificar.

La DIC ha demostrado, en general, todas estas características pero, como decíamos, probablemente es el único. La R-I fue sometida a un parecido análisis pero en principio los resultados experimentales no fueron positivos. En general, aunque veremos que existen otros problemas más específicos, la cuestión clave es que la R-I ha demostrado tener sólo influencia en el ámbito cognitivo y básicamente infantil. Pero mientras para unos ésta invalida la dimensión, y la coloca en el ámbito del desarrollo intelectual o madurativo, para otros no es así; simplemente hay que aceptar que la R-I no es estrictamente hablando un estilo cognitivo, sino un estilo de aprendizaje o, en todo caso, hay que aceptar distintos tipos de estilo cognitivo. Buela-Casal, De los Santos-Roig y Carretero-Dios (2000b) han realizado una excelente revisión sobre el tema que ahora destacaremos en sus elementos principales.

En primer lugar cabe destacar la postura de Riding y Cheema (1991), que afirman que las diferencias entre estilo cognitivo y de aprendizaje son evidentes. Para ellos el estilo cognitivo es el modo habitual de la persona para procesar información, esto es, para percibir, pensar, recordar, resolver problemas, etc.. Por su parte, el «*estilo de aprendizaje*», a pesar de surgir al amparo de los estilos cognitivos, no es un término sinónimo: hace referencia a como un individuo actúa para resolver problemas específicos, fundamentalmente tareas de aprendizaje o relacionadas con el ámbito educativo. La influencia de las estrategias de aprendizaje no es generalizable a situaciones de tipo social, afectivo, etc.. En resumen, el estilo cognitivo da idea de estructura, de visión personalista, mientras el estilo de aprendizaje da idea de proceso, de estrategia desarrollada para afrontar determinadas tareas. Naturalmente es posible ver a la DIC o a la R-I como estructura, como proceso o como ambas formas a la vez, aunque también es verdad que tal vez sólo se ajuste empíricamente a una de las posibilidades.

En este sentido el modelo de la cebolla de Curry (1983) propone integrar los conceptos de estilo cognitivo y estrategia o estilo de aprendizaje por niveles o estratos. Los niveles se van diferenciando en función de su mayor o menor dependencia de las contingencias ambientales. Así, el primer nivel (la capa más externa) se refiere a la preferencia instruccional, es decir, la preferencia del sujeto por un determinado ambiente de aprendizaje, una materia, una forma de aprender, etc.. Lógicamente esta preferencia es altamente modificable e inestable. En el segundo nivel están los estilos de procesamiento de la información; la forma en que el individuo tiende a percibir la información y a manejarla. Aunque el ambiente no influye directamente sobre estos estilos, sí puede hacerlo indirectamente a través de las estrategias de aprendizaje; éstas son modificables o más bien moldeables. Finalmente, en el nivel más interno, están los estilos cognitivos de personalidad que no sólo influyen en la manera en cómo se asimila y adapta la información sino en cómo el individuo tiende a responder en situaciones sociales, afectivas, etc. En este caso, como en los rasgos de personalidad, se asume que es difícil cambiar esta estructura genérica de funcionamiento de la persona. Buela-Casal y cols. (2000b) concluyen que el primer nivel de Curry haría referencia a las estrategias de aprendizaje, el segundo a la concepción de los estilos cognitivos como «procesos» y el tercero a los estilos cognitivos como «rasgos de personalidad».

Por su parte, el modelo de familias de Riding (1997; Riding y Cheema, 1991) asume que los distintos estilos cognitivos pueden ser agrupados en dos grandes «familias»: la global-analítica y la verbal-figurativa. La primera hace referencia a las diferencias encontradas en los modos de procesar la información: los que la dividen en sus partes componentes y los que prefieren una aproximación más holística. La segunda hace referencia a las diferencias encontradas en la representación de la información: los que prefieren trabajar con imágenes mentales y los que prefieren mantener una forma verbal o semántica.

Prácticamente la totalidad de los estilos cognitivos mencionados anteriormente, incluidos la DIC y la R-I, se agruparían en torno a la familia global-analítica. Intuitivamente, al revisar la conceptualización y las

pruebas de evaluación de estos estilos (véase Servera, 1997), es fácil tener la sensación de que prácticamente en todos un polo hace referencia a una aproximación analítica-secuencial de la información y el otro a una aproximación global-simultánea. De hecho, ello constituye una tradición en el estudio del funcionamiento intelectual de la persona que tiene un claro reflejo en la teoría neuropsicológica de Luria (donde en mayor o menor medida se propone esta misma diferenciación) o en la mayoría de las baterías más utilizadas para evaluar la inteligencia, como la K-ABC de Kaufman y Kaufman (1998). Por contra, y centrándonos en los datos que hacen referencia estrictamente a los estilos cognitivos, no ha sido tan fácil encontrar evidencias favorables a la constitución de esta familia. En primer lugar hay que centrarse casi exclusivamente a la relación de la DIC con la R-I (puesto que hay muy pocos datos con los otros estilos y además sus pruebas de evaluación psicométricamente son muy discutibles) y, aunque Buéla-Casal y cols. (2000b) son más optimistas en su revisión de los datos disponibles, la verdad es que la correlación entre las puntuaciones de las pruebas de la DIC y la R-I sólo son moderadas. Es verdad que existe una tendencia a que la reflexividad se asocie con la independencia de campo (o viceversa) pero observando detenidamente la naturaleza de las pruebas de evaluación de ambos estilos y asumiendo una raíz común en su forma de procesar la información cabría esperar correlaciones mucho más elevadas de lo que los trabajos de investigación hasta ahora han encontrado.

El modelo de familias de Riding es sin duda el modelo integrativo de los estilos cognitivos con más evidencias experimentales. Pero para algunos, como Jones (1997), todavía no son suficientes. Hay que clarificar, como mínimo, el papel de la dimensión reflexividad-impulsividad.

2.2.2 La dimensión reflexividad-impulsividad

A mediados de los años sesenta se produjo la conceptualización de la dimensión reflexividad-impulsividad (R-I) o, como en un principio se le denominó, el «*tempo conceptual*». Bornas y Servera (1996) han revisado en profundidad el surgimiento de esta dimensión, pero tanto antes

como después en nuestro país es posible encontrar otras revisiones relevantes, como la de Palacios (1982) o la de Buela-Casal, Carretero-Dios y De los Santos-Roig (2001). A continuación, de forma resumida, ofreceremos los aspectos más relevantes de dicha conceptualización.

Como comentábamos, a mediados de los sesenta, Jerome Kagan y otros investigadores llevaron a cabo en el Instituto Fels, vinculado al Centro de Estudios Cognitivos de la Universidad de Harvard, los trabajos que sirvieron para establecer la definición del estilo impulsivo, su evaluación y sus implicaciones. En un principio la línea de trabajo del grupo de Kagan tenía como objetivo el estudio de los llamados estilos de conceptualización. La idea de la que partían es que existen diferencias interindividuales e intraindividuales en la estructura y el contenido de las categorías conceptuales que realizan las personas, producidas cuando la situación es relativamente libre y cuando existe una cierta variedad de dimensiones que pueden servir de base para realizar las categorizaciones. Habitualmente los estilos de conceptualización se medían presentando un conjunto de dibujos de personas, animales u objetos a una muestra de sujetos para que éstos los agruparan basándose en algún criterio que después era convenientemente categorizado. Kagan y sus colaboradores habían hipotetizado que los adultos que previamente se habían mostrado como «*más inteligentes*» preferirían hacer sus agrupamientos sobre la base de criterios más abstractos o categoriales (por ejemplo, jóvenes-viejos, enfermos-sanos, etc.) pero los datos experimentales no confirmaron esta hipótesis (Kagan, Moss y Sigel, 1963). La realidad demostró que las personas más inteligentes, más independientes y más atraídas por las actividades intelectuales preferían realizar sus agrupamientos basándose en un elemento objetivo compartido por todos los dibujos (por ejemplo, juntaban todas aquellas figuras que llevaban un objeto en la mano). Este tipo de agrupamientos eran denominados «*analíticos*» porque implicaban tomar en consideración un componente específico a partir de un conjunto relativamente amplio de estímulos. A parte de la agrupación categorial y ésta analítica, también es posible distinguir clasificaciones funcionales, ya sean relacionales (la base del agrupamiento implica una relación te-

mática —cerillas, pipa—) o localizadas (los miembros comparten una localización común —animales que viven en una granja—).

El trabajo con los tests de estilos conceptuales se realizó intensamente comparando adultos con niños y siguiendo a éstos longitudinalmente. Ello sirvió para concluir, entre otras cosas, que las clasificaciones analíticas, propias de los sujetos más «*inteligentes*» y más deseosos de «*aprobación social*», aumentan progresiva y linealmente con la edad, concretamente la evolución que siguen es la siguiente: al principio predominan las agrupaciones funcionales, luego se pasa a las categoriales y al final se llega a las analíticas. Por otra parte, lo que en principio pareció coyuntural después resultó definitivo: la producción de conceptos analíticos se asocia con una tendencia a retrasar la respuesta, es decir, con un tiempo de latencia más largo. La repetida constatación de esta última asociación despertó el interés de otros investigadores. Pronto se descubrió que con la edad y con el aumento de conceptualizaciones analíticas también aumentaba esta latencia de respuesta; pero lo más sorprendente fue descubrir que los niños «*más analíticos*» no sólo empleaban más tiempo, sino que cometían muchos menos errores a la hora de resolver tareas de emparejamiento de objetos similares y, además, presentaban cierta superioridad en capacidades cognitivas y muy especialmente en el rendimiento académico (Kagan y cols., 1964).

Estos datos que, como comentábamos, en principio eran relativamente colaterales a los objetivos de la investigación en estilos conceptuales poco a poco fueron ocupando el primer lugar de interés. Kagan y Kogan (1970), en su amplia revisión, explicaron las causas del paso de los estilos conceptuales a los estilos cognitivos, con la posterior polémica que está generó, y que comentamos anteriormente. El objeto de estudio se fue centrando en la postura impulsiva o reflexiva ante distintas tareas de laboratorio y de rendimiento, y cada vez más se asentó la visión bipolar de la dimensión. El trabajo siguió, más como expresión de la curiosidad de los investigadores que como línea perfectamente definida. Kagan, Moss y Siegel (1963, p. 204) ya reconocían en los albores de esta dimensión que en su trabajo había mucho de «*ensayo y error*». Palacios (1982) corrobora esta afirmación y pone el énfasis en la gran

cantidad de pruebas dispares que desde un principio son creadas para observar el fenómeno del retraso de la latencia de respuesta. En la elaboración de estas pruebas la única premisa a respetar era que el sujeto siempre debía elegir una solución entre varias alternativas posibles, de modo que aumentaba mucho la posibilidad de elicitar un tipo de respuesta impulsiva frente a otra reflexiva.

Tal vez por la propia inercia del método del «ensayo y error», o tal vez por las múltiples publicaciones que en muy pocos años aparecen en referencia a la dimensión reflexividad-impulsividad cuando ésta aún no estaba perfectamente definida, se da desde el principio un cierto desconcierto en la forma en que deben entenderse los estilos reflexivo e impulsivo, que perduró durante muchos años y aún hasta hoy. Gilmore (1968) fue de los primeros en señalar que muchos investigadores y profesionales de la época entendían que los niños reflexivos eran aquéllos que tendían a retrasar su respuesta ante la mayoría de situaciones estímulares, es decir, como si la latencia de respuesta pudiese ser conceptualizada en forma de rasgo de personalidad. En realidad esta idea no se corresponde a la fundamentación del trabajo del grupo de Kagan, aunque hay que reconocer que algunas de sus definiciones del término «tempo conceptual» podían inducir a tal confusión:

«Hemos acuñado la expresión "tempo conceptual" para describir el significado connotativo que asignamos a la variable reflexividad. Algunos niños dan rienda suelta de manera consistente a la primera hipótesis razonable que se les ocurre sin pararse a reflexionar sobre su posible validez. Su estrategia de resolución de problemas es como un arma de fuego: el niño dispara una descarga de respuestas con la esperanza de que una de ellas sea correcta, o, tal vez, porque necesita un feedback inmediato del entorno que le informe de la calidad de su elección. Este niño contrasta con el que de forma típica se detiene a considerar la validez diferencial de varias hipótesis. Este niño actúa como si deseara fuertemente ser tan correcto como es capaz en su primer intento y pudiese tolerar la ambigüedad y tensión inherentes al período de silencio que es un concomitante inevitable de la selección de la respuesta» (Kagan, 1966a, pp. 502-503, citado en Palacios, 1982 p. 33).

La verdad es que aunque la anterior definición pondera básicamente el tiempo de respuesta como la variable clave del funcionamiento cognitivo que se estudia, está ligeramente sacada de contexto. En primer lugar el contexto estimular en donde se da la reflexividad-impulsividad no es tan laxo: desde los primeros trabajos (Kagan y cols., 1964; Kagan, 1965a; Kagan, 1965d; Kagan, 1966a; Kagan 1966b) se recalca el hecho de que la dimensión es propia de las situaciones problema que contienen un cierto grado de incertidumbre de respuesta, es decir, situaciones en donde aparecen simultáneamente varias soluciones pero sólo una de ellas es la correcta. Con ello queda claro que en otro tipo de tareas, por ejemplo, las que se basan en conocimientos previos o en razonamientos lógico, o en cualquier otra donde la incertidumbre de respuesta no es un factor clave la dimensión no tiene porqué mantener el mismo nivel de influencia. En segundo lugar, Kagan y cols. (1964) también postulan que una aproximación reflexiva a la resolución de un problema también debe asociarse a una probabilidad menor de cometer errores, o a la inversa para el caso de la impulsividad. Es por eso que en el mismo trabajo que anteriormente hemos citado, sólo unas páginas más adelante de donde hemos entresacado la anterior definición, también se expone claramente que el adjetivo «*reflexivo*» es «*sobre todo descriptivo del niño que presenta largas latencias y pocos errores*» (Kagan, 1966a, p. 490).

En definitiva, podemos definir el estilo cognitivo reflexividad-impulsividad (R-I) como un índice de la capacidad analítica de las personas que se define a partir de las puntuaciones de errores y latencias en tareas de discriminación visual que implican incertidumbre de respuesta. Aunque parece una definición suficientemente operativizada necesariamente sigue manteniendo problemas conceptuales: ¿por qué mantener el término «*estilo cognitivo*» si ya hemos visto que en realidad parece no cumplir del todo con los requisitos de este constructo teórico? y ¿por qué seguir hablando de estilo cognitivo o de aprendizaje cuando introducimos la variable errores en su evaluación lo que nos lleva directamente al tema de la capacidad o «*inteligencia*»? Vamos a intentar aclarar estas dos cuestiones porque en definitiva resultan básicas para mantener el interés de investigadores y profesionales en su evaluación.

Ciertamente, como ya hemos comentado, denominar «*estilo cognitivo*» en sentido estricto a la R-I presenta muchas dificultades. Jones (1997) y Servera (1997) destacan los principales problemas que dificultan la conceptualización de la dimensión como estilo cognitivo (especialmente en comparación con la dependencia-independencia de campo, DIC):

- La R-I ha mostrado influencia especialmente en el ámbito del aprendizaje, pero en mucho menor grado en el comportamental, social o afectivo.
- Mientras la DIC y otros estilos cognitivos mantienen su interés en la edad adulta, muy pocos trabajos han encontrado relevante la diferenciación entre impulsivos y reflexivos más allá de los 12 años, incluso las pruebas de evaluación por encima de este límite son escasas y discutibles.
- Mientras en principio procesar globalmente o analíticamente no debe implicar juicios de valor, presentar un estilo «*impulsivo*» (en teoría holístico más que analítico) sí suele asociarse con más dificultades de aprendizaje, presentando una connotación negativa.
- Los estilos cognitivos, al menos entendidos en la vía personalista, ni son modificables, ni tienen porqué ser modificados, más bien los efectos que puedan tener dependerá de las características estimulares de la tarea o la situación. Pero como acabamos de comentar, la connotación negativa de la impulsividad ha obligado a buscar formas de modificación de esta tendencia, con resultados moderados pero aceptables.

Todos estos argumentos llevaron a Jones (1997) a considerar a la R-I más como una estrategia de aprendizaje que como un estilo cognitivo en el sentido tradicional del término. Aunque estamos de acuerdo esencialmente con esta propuesta, hay que matizar que no todo el mundo entiende lo mismo por «*estrategia de aprendizaje*». Por ejemplo, a veces las técnicas o habilidades del subrayado, de la regla de tres o de elaboración de esquemas se presentan como «*estrategias de aprendizaje*» cuando más bien son sólo eso, habilidades, o en términos de Nisbet y

Shucksmith (1987), funciones «no ejecutivas», mientras las estrategias de aprendizaje son funciones ejecutivas, de planificación, control y acción del pensamiento. En este marco, o en los términos de redefinición de estilos cognitivos como estilos de procesamiento de información tal y como hacen los modelos de instrucción en estrategias cognitivas (véase McCormick, Miller y Pressley, 1989), es donde, en nuestra opinión, mejor se entiende la dimensión reflexividad-impulsividad. Pero retomaremos esta idea al final del presente apartado, cuando entremos más de lleno en estos modelos y la reconceptualización de la impulsividad. Ahora hemos dejado pendiente la cuestión crucial que relaciona la R-I con el campo de la inteligencia.

Como hemos visto, la definición de la dimensión R-I se hace en función del tiempo de respuesta del sujeto, pero también de su número de errores en tareas de discriminación que implican incertidumbre de respuesta. Para muchos autores, especialmente el llamado grupo de Block, la inclusión de una puntuación de capacidad desvirtuó totalmente la posibilidad de considerar a la R-I una dimensión estilística. De hecho, como veremos en la parte dedicada a la validez de constructo, ciertamente son los errores, y no las latencias, las que presentan las correlaciones más elevadas y significativas con cualesquiera sean las variables criterio que se pretendan evaluar: de rendimiento o de comportamiento del niño. Tal vez el primer punto crítico a destacar de esta conclusión es que la R-I no debe entenderse como el resultado de dos puntuaciones por separado, sino que errores y latencias deben constituir un único índice, pero éste sería otro problema. Ahora nos interesa destacar por qué la R-I no es un «artefacto» de la capacidad intelectual del niño, y aún estando relacionado con ella mantiene una respetable independencia.

La revista *Child Development* durante los años setenta y ochenta fue testigo de excepción del enfrentamiento entre Jack Block y sus colaboradores de la Universidad de Berkeley y el grupo de Kagan. Block, Block y Harrington (1974) presentaron un amplio estudio con niños entre 4 y 5 años utilizando el «*Matching Familiar Figures Test*» (MFF) de Kagan, el test Q de personalidad que cumplimentaron maestros, y el WPSSI para conseguir el CI de los niños, para concluir, entre otras cosas, que la contribución

de las latencias a la variables de comportamiento y de inteligencia es prácticamente nula, y que la puntuación de errores era explicable por el CI. Kagan y Messer (1975) replicaron estas afirmaciones, tanto por el hecho de que ellos no entendían que la R-I debiera proyectarse sobre tests de personalidad de origen psicodinámico, como por utilizar niños de preescolar. Block, Block y Harrington (1975) argumentaron la validez de sus conclusiones apelando a otros trabajos con niños de enseñanza primaria con similares conclusiones: los errores eran la variable clave, y casi siempre correlacionaban de modo significativo o con el CI (Ault, 1973; Denney, 1973) o con la DIC (Campbell y Douglas, 1972). En 1986 Block, Gjerde y Block replicaron su propio estudio original, ahora con niños entre 11 y 14 años, utilizando prácticamente las mismas pruebas (ahora el WISC era la prueba de inteligencia) y de nuevo observaron como las latencias, una vez controlado el efecto de los errores, prácticamente no correlacionan con ninguna variable, y todas las correlaciones significativas de los errores vienen limitadas por una elevada correlación con el CI, de entre $-.51$ y $-.61$. En resumen, según estos autores, qué necesidad tenemos de complicar las diferencias encontradas sobre el MFF atribuyéndolas a estrategias de procesamiento de la información, cuando desde la óptica del rendimiento intelectual las podemos explicar satisfactoriamente.

La postura del grupo de Block, sin embargo, no siempre ha tenido tanto apoyo empírico como el que se deduce de sus propios datos. En una de las primeras revisiones sobre la R-I, Messer (1976) incluye 23 estudios que han utilizado algunos de los tests clásicos de medida de la inteligencia, como el el Stanford-Binet, el test de Aptitudes Mentales Primarias (PMA), y las escalas de Wechsler (especialmente el WISC), el Kuhlman-Anderson, el Lorge-Thorndike, el Otis-Lennon, etc. Las conclusiones de los datos revisados se pueden resumir así:

- La correlación media entre las latencias del MFF y el CI es de $.14$ para los niños y algo superior para las niñas, $.22$, con un índice medio del $.17$.
- La correlación media entre los errores del MFF y el CI es, para los niños, de $-.30$ y, para las niñas, de $-.34$. Por consiguiente, en prin-

cipio podemos considerar que entre la R-I y la inteligencia existe una relación moderada, si bien atribuible fundamentalmente a los errores.

- Si dividimos los test entre los de contenido verbal y los no verbales podemos observar que estos últimos correlacionan más con los errores y las latencias del MFF. Pero si además entre éstos seleccionamos los que obligan a dar una respuesta concreta y los que permiten elegir una respuesta a partir de varias alternativas, entonces se observa que cuanto más incertidumbre de respuesta contiene un test más correlacionará con el MFF.

En resumen, cabe hablar de una relación moderada especialmente relacionada con tests de inteligencia no verbales que implican incertidumbre de respuesta. En este sentido, Larsen (1982) trabajando con niños en edad escolar encontró que la correlación entre los errores del MFF y la puntuación de la serie II de las Matrices Progresivas Avanzadas de Raven fue de .22, y en el caso de las latencias de -.23. En nuestro caso con una muestra de 81 niños entre 10 y 11 años encontramos resultados algo divergentes: la correlación con los errores fue bastante superior, del -.38, mientras la de las latencias fue prácticamente insignificante, del .09 (Servera, 1992). Como indica Wiedl (1980), aún sin ofrecer índices de correlación existen otros muchos estudios en donde se ha podido constatar que los impulsivos puntúan más bajo que los reflexivos sobre el Raven. Pero el Raven realmente es una prueba de inteligencia que aún basándose en la capacidad de razonamiento abstracto también implica un factor importante de incertidumbre de respuesta, al igual que las medidas de R-I.

En definitiva, de los trabajos revisados y otros muchos que han trabajado el tema (Gow y Ward, 1981, 1982a, 1982b; Egeland, Bielke y Kendall, 1980; Genser, Häefele y Häefele, 1978) podemos concluir que, aún considerando una importante variabilidad atribuible a las características del tipo de test, el rendimiento de los reflexivos siempre acostumbra a ser ligeramente superior al de los impulsivos. Ahora bien, esta superioridad raramente llega a ser estadísticamente significativa. Por otra parte, la su-

perioridad de los reflexivos sobre los impulsivos, con independencia de que no siempre llegue a ser relevante, es prácticamente demostrable en todas las tareas de rendimiento cognitivo (especialmente las académicas, pero también otras de laboratorio). La explicación a esta relación ya la esbozó acertadamente Nathan Kogan en 1976; las pruebas de evaluación intentar elicitar distintas formas de afrontar una tarea de discriminación con incertidumbre de respuesta. Los tests de inteligencia no son una excepción: en general cuando está presente la necesidad de considerar distintas opciones en un grado de incertidumbre elevado la relación de la inteligencia con la reflexividad aumenta claramente, en caso contrario se atenúa. ¿Puede explicar la inteligencia, el factor G, los datos que dependen de la R-I? Probablemente no: ese factor general de inteligencia también está presente en las pruebas que evalúan otros estilos cognitivos, como la DIC, o en cualquier otra actividad intelectual pero sería absurdo pretender reducirlo todo a su efecto. La correlación con la medida de la R-I es moderada, como con otras muchas medidas, pero no llega a ser invalidante. Por otra parte, como comentábamos, en la práctica totalidad de estudios donde han aparecido resultados sobre tests de inteligencia, aún mostrándose superiores los reflexivos, no siempre ha sido de modo significativo y prácticamente nunca los impulsivos han estado por debajo de los intervalos de normalidad.

Retomamos ahora la cuestión de la relación entre la R-I y la DIC, que como ya comentamos en el apartado dedicado a los estilos cognitivos y de aprendizaje también aparece difusa. Para validar el modelo de familias de estilos cognitivos de Riding cabría esperar que la relación fuese elevada, y en el sentido de mayor reflexividad a mayor independencia de campo (factor analítico) o viceversa, mayor impulsividad a mayor dependencia (factor globalista). Pero, en nuestra opinión, sólo cabe hablar, como en el caso de la inteligencia, de una relación moderadamente elevada, eso sí, en la dirección esperada. En la ya citada revisión de Messer (1976) también se analizan bastantes trabajos que intentaron establecer desde el principio la relación entre la R-I y la DIC con muestras de preescolar y de edad escolar. En principio los coeficientes extraídos deberían restringirse a la puntuación continua del EFT (la medida clásica-

ca de la DIC) con los errores y las latencias del MFF, pero algunos son más completos y también midieron la latencia de respuesta sobre el EFT. Los resultados sólo son coincidentes en un aspecto: todos los estudios refieren correlaciones moderadas entre los errores del MFF y la puntuación de las distintas versiones del EFT utilizadas (con un rango que va del $-.35$ al $-.62$). Sólo dos estudios encuentran correlaciones alrededor del $.40$ entre las latencias del MFF y del EFT, mientras que en un solo estudio aparece una ligera correlación ($-.17$) entre los errores del MFF y las latencias del EFT. Es decir, que la moderada correlación que se esboza entre ambos estilos cognitivos en todo caso vendría determinada casi exclusivamente por la puntuación de errores. En trabajos posteriores se ha confirmado plenamente esta tendencia, si bien, como ya señaló acertadamente Kogan (1976), si los autores hubiesen sido más cuidadosos a la hora de controlar la variable «edad» en los estudios longitudinales es muy probable que los coeficientes fuesen aún más bajos.

En nuestros propios trabajos (Servera, 1990) también obtuvimos algunos datos al respecto de esta cuestión. El estudio incluyó una muestra de 47 niños de preescolar, 103 de los primeros cuatro cursos de primaria y 94 de los últimos tres cursos (de 10 a 12 años). Para cada grupo de edad se utilizó la prueba de evaluación de la DIC adecuada y para todos el MFF20 (una versión mejorada del MFF clásico). Los resultados destacaron la nula presencia de correlaciones en preescolar. En los primeros cursos de primaria la correlación de los errores del MFF20 con la puntuación de la DIC fue del $-.45$ y la de las latencias del $.23$. Finalmente en los últimos cursos de primaria la correlación fue del $.40$ en los errores (aunque el signo de la correlación aparezca invertido siempre es en la misma dirección, a menor puntuación en errores, mayor independencia de campo y viceversa) y desapareció en el caso de las latencias. Estos resultados no discrepan en exceso de los revisados por Messer (1976), exceptuando la nula relación que se observa en preescolar, debida, quizás, a la baja consistencia interna de nuestra prueba abreviada del MFF20. En definitiva, tal y como era de esperar, se confirma que existe una tendencia a que los niños que presentan puntuaciones hacia la reflexividad, también las tengan hacia la independencia de campo,

pero se trata de una tendencia moderada y atribuible en gran parte a la puntuación de errores.

Existen básicamente dos formas de entender la naturaleza de esta moderada relación, en absoluto excluyentes. La primera se fundamenta la presunción de que las pruebas de evaluación de la DIC y la R-I comparten el carácter perceptivo y discriminativo, pero en las de la DIC no está presente en grado tan evidente la incertidumbre de respuesta, luego al faltar un elemento clave elicitor de la estrategia R-I es lógico que la correlación no sea muy elevada. La segunda, en nuestra opinión más interesante, deriva de lo que ya postuló Neimark (1975) hace veinticinco años: aunque exista un tipo de procesamiento de la información preferido y relativamente estable la inmensa mayoría de sujetos dispone de recursos para poner en marcha distintas estrategias de actuación; por tanto, probablemente lo que caracterice a los sujetos con mejor rendimiento intelectual es su mayor capacidad de flexibilización de los recursos cognitivos y no el disponer de unos en mayor grado que otros. Desde este punto de vista, los índices de correlación entre la DIC y la R-I serían atribuibles básicamente al grupo de sujetos de «*superior flexibilidad*» que no es sólo que sean más reflexivos o más independientes de campos, sino que además probablemente sean más analíticos y con mayor capacidad de solución de problemas en general.

Retomamos ahora, para acabar este apartado, la idea anteriormente esbozada en referencia a la reconceptualización de la impulsividad, y del propio sentido de los «*estilos cognitivos*», en los llamados modelos de instrucción en estrategias cognitivas. En esos modelos nosotros incluimos, entre otros, el modelo de «*solución de problemas*» de Baron (1985), el modelo «*componencial*» de Sternberg (1982) (luego asimilado a su teoría triarquica de la inteligencia) y el modelo del «*usuario de buenas estrategias*» de Michael Pressley (1986). Estos modelos, en su momento, pero aún hoy en día en sus múltiples reformulaciones, han tenido gran influencia en la concepción del enfoque educativo del «*enseñar a pensar*», en el énfasis en la necesidad de enseñar estrategias más que contenidos y, en definitiva, en el objetivo de formar personas «*reflexivas*», con conocimiento metacognitivo, flexibles y con capacidad de adaptación a un

mundo donde la información cambia por rápidamente. En todos esos modelos, quizás preferentemente en el de Pressley, la impulsividad, o lo que ellos mismos denominan el «*estilo cognitivo impulsivo*» juega un papel muy relevante. Para ellos el término *estilo cognitivo* pierde gran parte de su connotación personalista, y pasa a denotar más bien *estilo preferido* de procesar y manejar la información. Así, por ejemplo, en el modelo de Pressley la resolución de problemas implica los siguientes aspectos: conocimiento y secuenciación de estrategias de actuación (estrategias más o menos generales), conocimiento metacognitivo de las estrategias de actuación (cómo y cuándo utilizar las estrategias), conocimiento de base o declarativo (memorístico, habilidades de aprendizaje, contenidos, etc.), *estilo atribucional* (autoestima, autoconfianza, aspectos motivacionales, etc.) y, finalmente, *estilo cognitivo*, destacándose la incompatibilidad de hacer funcionar una buena heurística de solución de problemas con una aproximación impulsiva, de bajo control.

Sin duda en este ámbito, el de sesgo en la forma de procesar la información y afrontar las tareas de aprendizaje, es donde la R-I, a nuestro juicio, alcanza su máxima relevancia. No hay que pretender, pues, que los niños de *estilo impulsivo* demuestren diferencias en su comportamiento social, afectivo, o incluso intelectual de un modo genérico, sino que su característica definitoria está su forma de aproximarse a tareas de solución de problemas, y de aprendizaje en general. Dadas las características de nuestro sistema educativo, y del tipo de tarea que solemos emplear, probablemente el rendimiento de estos niños será inferior a lo deseado. Pero ello no será atribuible a un problema de falta de conocimientos o de capacidad, sino a la propia forma de afrontar las tareas. Su *estilo* dificulta el aprendizaje secuenciado, planificado y ordenado.

2.3 La evaluación del estilo impulsivo: los tests de emparejamiento de figuras familiares (MFF)

En el trabajo de ensayo-error que hemos ido definiendo en la conceptualización de la R-I también hay que incluir las distintas pruebas de eva-

luación, prácticamente todas de laboratorio. Sin embargo, desde los primeros tiempos destacó una, que hoy en día es muy popular, el Matching Familiar Figures Test (MFF) o test de emparejamiento de figuras familiares (Kagan, 1965c). A lo largo de las más de tres décadas y un lustro esta prueba no ha dejado de utilizarse en el ámbito escolar y clínico (especialmente en casos de hiperactividad infantil), sin que tampoco sus detractores hayan dejado de criticar aspectos conceptuales y metodológicos, y otros autores, entre los que nos encontramos nosotros con el presente trabajo, tampoco han dejado de buscar nuevas versiones más potentes.

El MFF se define como una prueba de emparejamiento perceptivo en el que no intervienen de manera significativa ni la memoria ni el razonamiento, aplicable a niños entre 6 y 12 años de edad aproximadamente. La prueba original consta de 12 ítems con idéntica estructura: el sujeto debe observar un modelo (un dibujo de una persona, animal u objeto familiar) y seis figuras alternativas muy parecidas, de entre las cuales debe hallar la que es exactamente igual al modelo, teniendo en cuenta que sólo hay una respuesta acertada y que no tiene un límite de tiempo para hallarla. La forma de aplicación del test es individual y las instrucciones son muy sencillas. Para su aplicación basta que el niño tenga claro el significado de los conceptos «igual» y «diferente». El test contiene dos ítems de práctica que permitirán al experimentador cerciorarse de que el sujeto ha comprendido perfectamente en qué consiste la tarea. Deben tenerse en cuenta dos cosas muy importantes: la primera es que el niño debe tener siempre a la vista el modelo y las alternativas, y la segunda es que el experimentador debe informarle después de cada elección de si es o no correcta.

El experimentador debe registrar dos tipos de respuestas sobre cada ítem del MFF. La primera es el tiempo, en segundos, que tarda el niño en dar la primera respuesta, sea o no acertada. Y la segunda es el número de errores que comete en cada ítem. Al final se define una puntuación de errores: número total de errores a lo largo del test, y otra de latencias: tiempo medio medio de respuesta para el total de ítems.

La interpretación de las puntuaciones del MFF, como ya hemos visto conceptualmente en el anterior apartado, genera bastantes controver-

sias. Sin duda el problema de base está en conjugar dos puntuaciones teóricamente de naturaleza muy distinta: la latencia de respuesta (que sería la estilística pura, aunque ya hemos visto que no es definitiva por sí sola de la R-I) y la de errores, que es de competencia, de capacidad. Tras diversos intentos el grupo de Kagan llegó a la conclusión que, al contrario de otros tests, resultaba muy difícil establecer una baremación en función de estas dos puntuaciones y, por otra parte, tampoco la idea de evaluar el estilo R-I iba muy en consonancia con la corriente psicométrica, sino más bien con la experimental. En este sentido se propuso que múltiples factores sociales, culturales o personales podrían influir a la dimensión, de manera que podría tener poco sentido buscar tablas de baremación en las que ubicar a cualquier sujeto independientemente de estos factores. Todo ello desembocó en la propuesta de clasificación de los sujetos denominada de la «doble división por las medianas de errores y latencias», también conocida como clasificación «tradicional» o por el sistema de «cuadrantes».

El sistema tradicional se desarrolla de la siguiente manera: en primer lugar se aplica el MFF a una muestra cualquiera de niños de edades similares (por ejemplo, los que pueden componer una aula de cualquier curso escolar). En segundo lugar se calcula la mediana (es decir, la puntuación que asegura el dejar por arriba y por debajo el mismo número de sujetos) tanto de la puntuación de error como de latencias, de manera que se obtiene una distribución por cuadrantes. Y, por último, cada niño es asignado a uno de los cuadrantes en función de de las normas que aparecen en el siguiente esquema:

ERRORES	CUADRANTES DE ESTILO VS. CAPACIDAD	
Por encima de la mediana	IMPULSIVOS	LENTOS-INEXACTOS
Por debajo de la mediana	RÁPIDOS-EXACTOS	REFLEXIVOS
	Por debajo de la mediana de latencias	Por encima de la mediana de latencias
LATENCIAS		

Como puede observarse en dicha tabla, el resultado de la aplicación del sistema obliga a aceptar la presencia de estilos de respuesta versus niveles de capacidad de respuesta. Ciertamente los niños que interesan para el estudio del estilo R-I son los «impulsivos» y los «reflexivos», pero siempre es inevitable el tener que considerar la presencia de sujetos rápidos-exactos y lentos-inexactos, es decir, aquéllos cuya velocidad de respuesta (corta o larga) no parece afectar al número de errores y, por tanto, su ejecución parece mucho más ligada a un factor de competencia o incluso de CI. La presencia de estos niños se considera una especie de «artefacto», de hecho su número depende en gran medida del grado de correlación negativa entre los errores y las latencias: cuanto mayor sea dicha correlación menor será la presencia de sujetos rápidos-exactos y lentos-inexactos. Como veremos más adelante, al abordar la validez de constructo de la R-I, dicha correlación es un factor clave, y el primer dato a examinar en sus pruebas de evaluación.

En el caso del MFF los resultados parecen señalar que entre los 6 y 12 años dicha correlación está alrededor del $-.50$, de todas formas, la mayoría de las investigaciones coinciden en destacar su gran variabilidad (Haskins y McKinney, 1976; Salkind y Wright, 1977; Mitchell y Ault, 1979; Pratt y Wickens, 1983). En cualquier caso, sea como fuere, lo que es evidente es que con este sistema la comparación entre sujetos, fuera de su muestra de referencia, es muy complicada. Como señala Palacios (1982, p.37), «la clasificación de un sujeto reflexivo o impulsivo es siempre relativo a los datos del grupo en el que se incluye, por lo que, por poner un ejemplo, un sujeto que estuviera hipotéticamente formando parte de dos muestras podría ser considerado en una de ellas como reflexivo y en la otra como impulsivo». No obstante ha habido algunos intentos para conseguir puntuaciones normalizadas del test MFF. Margolis y Branningan (1976) presentaron algunos datos normativos aunque sólo pueden ser tomados en cuenta, según los propios autores, trabajando con muestras del área metropolitana de Nueva York. En el trabajo de revisión de Messer (1976) se presenta un primer intento de ofrecer datos normativos a partir de la síntesis de las principales investigaciones llevadas a cabo hasta la fecha, aunque tiene importantes lagunas. Neil J.

Salkind (1977, 1978) ha publicado normas para las edades comprendidas entre los 5 y los 12 años trabajando con una muestra de 2.800 sujetos. De todos modos, a excepción de los colaboradores de Salkind, la verdad es que ha habido pocos autores que hayan tenido en cuenta estas puntuaciones a la hora de clasificar a los sujetos como impulsivos y reflexivos con fines científicos; básicamente sus puntuaciones han tendido una finalidad clínico-profesional.

Por otra parte, en otro interesante trabajo Salkind y Nelson (1980) presentan la evolución de las puntuaciones de error y latencias sobre el test MFF de los 5 a los 12 años. Los autores observaron que con la edad la puntuación de errores desciende de forma pronunciada y constante hasta su estabilización alrededor de los 10-11 años. Por su parte, la puntuación de latencias experimenta un aumento constante, entrecruzándose con la de errores sobre los 8 años, y estabilizándose alrededor de los 10 años. Por tanto, según los datos de Salkind y Nelson (1980) las puntuaciones de error y latencias siguen una evolución inversamente proporcional hasta los 10-11, que es la edad donde se estabilizan y donde el test muestra ya el efecto de «techo». En cualquier caso el MFF ha venido arrastrando problemas en sus índices psicométricos que han alentado a los investigadores a desarrollar versiones más específicas o mejoradas.

El propio grupo de Kagan desarrolló tres formas del MFF: la K para preescolares (muy poco utilizada), la F y la S como versiones paralelas para la edad escolar. Por su parte, Banta (1970) creó un test muy poco utilizado el EC-MFF (Early Childhood Matching Familiar Figures Test) que únicamente emplea tres alternativas para cada modelo. Por otra parte, y aunque el propio Kagan (1981) lo desestimó como una buena medida de la R-I, el test más utilizado en edad preescolar es el KRISP (Kansas Reflection-Impulsivity Scale for Preschoolers) de Wright (1971, 1973). Desafortunadamente en un trabajo propio también pudimos comprobar que no es una medida muy fiable (Bornas, Servera y Montaña, 1998). Existen otras múltiples versiones, algunas para adultos (Yando y Kagan, 1968) y otras que han alterado el sistema de presentación de los ítems, como el MFF Sequential Presentation Test» (Butter y

Jeffcott, 1982; Butter y Snyder, 1982), pero realmente la versión de referencia, la que ha presentado los mejores resultados es el MFF20 de Cairns y Cammock (1978).

Los trabajos de Cairns y Cammock (1978) en la Universidad del Ulster sirvieron para confirmar algunas críticas psicométricas que Egeland y Weinberg (1976) y Ault, Mitchell y Hartman (1976) ya habían hecho en referencia al MFF: la necesaria correlación entre errores y latencias era bastante menos alta de lo que se presumía, e igualmente eran bajos los índices de fiabilidad por consistencia interna o por test-retest. Los autores decidieron elaborar un nuevo test trabajando a partir de la mejora de la capacidad discriminativa de alguno de los mejores ítems del MFF y elaborando otros de parecidas características. Al final de varios estudios psicométricos surgió el MFF20; un test de 20 ítems que consta de dos ítems de prueba y que sigue las mismas normas de aplicación ya especificadas para el MFF. En nuestro país, que nosotros tengamos conocimiento, el test MFF20 fue dado a conocer en la revisión de Palacios (1982) y fue utilizado inicialmente, al menos, en los trabajos de Navarro (1987), Quiroga y Forteza (1988) y en nuestras propias investigaciones (Servera, 1990; Bornas y cols., 1990; Servera, 1992). Últimamente un grupo de Granada, encabezado por Gualbuerto Buela-Casal, ha desarrollado una intensa labor de investigación sobre esta prueba, realizando una adaptación española (Buela-Casal, Carretero-Dios, De los Santos-Roig y Bermúdez, en prensa).

2.3.1 Aspectos psicométricos de los MFF

Vamos a revisar a continuación básicamente aspectos de fiabilidad y validez del MFF y del MFF20. Habitualmente la fiabilidad de la mayoría de tests admite ser estudiada tanto a través de su consistencia interna como a través del grado de estabilidad de sus puntuaciones. Normalmente en el caso de la consistencia interna se utiliza el método de la división por las dos mitades o el cálculo del coeficiente alfa de Cronbach, mientras en los estudios de estabilidad se opta por los métodos de test-retest o por el de «*formas paralelas*».

Por lo que respecta a la fiabilidad del test MFF, en el caso de los estudios de la estabilidad de las puntuaciones ya podemos adelantar que al utilizar los procedimientos antes mencionados se han obtenido índices de fiabilidad bastantes bajos, aunque, como han señalado Gjerde, Block y Block (1985), en el caso del MFF la ausencia de estabilidad no tiene porqué indicar necesariamente baja fiabilidad ya que se han dado una serie de circunstancias particulares: en primer lugar no se han construido «*formas paralelas*» que presentasen garantías de serlo y, en segundo lugar, no se han podido establecer intervalos «*adecuados*» del intervalo test-retest. A todo ello hay que añadir, como también señalan Gjerde y cols. (1985), por un lado la circunstancia especial de este test que supone establecer el nivel de fiabilidad para dos puntuaciones de distinta naturaleza y, por otro lado, la gran disparidad en el tipo de muestra en los estudios publicados. Con todos estos condicionantes Bornas y Servera (1996, p. 34 y ss.) analizaron algunas de las principales revisiones sobre el tema (Messer, 1976; Cruz, 1987 y Navarro, 1987) y concluyeron que tomando los resultados en su globalidad podríamos decir que el índice de fiabilidad fluctúa, tanto para las latencias como para los errores, entre el .45 y el .55. Es decir, unos índices de fiabilidad bastante bajos, si bien sujetos a variabilidad en función del intervalo en las aplicaciones test-retest, en el tipo de muestra, etc.

Por lo que respecta al cálculo de la fiabilidad a través de los procedimientos de consistencia interna, las conclusiones también se ven afectadas por la disparidad de los resultados, aunque en general se observa una clara tendencia a que las latencias manifiesten una mayor consistencia interna que los errores: el coeficiente de las latencias oscila entre el .86 y el .95, mientras por contra el coeficiente de los errores oscila entre el .47 y el .78. Es decir, mientras la fiabilidad estimada de las latencias sería alta, la de los errores en muchos casos parece ser demasiado baja.

Por lo que respecta a los estudios de consistencia temporal, es decir, la capacidad del test para mantener una relativa ordenación de los sujetos dentro de su propio grupo o cohorte a través del tiempo, el más conocido es el de Gjerde y cols. (1985). Estos autores realizaron un es-

tudio longitudinal donde testaron en cuatro ocasiones (a los tres, cuatro, cinco y once años) a una muestra compuesta por 64 niños y 64 niñas. En total, para cada sexo, realizaron seis correlaciones temporales para las latencias y seis para los errores. Los resultados indican que las cuatro correlaciones temporales de los errores estadísticamente significativas (tres para los niños y una para las niñas): .44, .50, .55 y .56, con la corrección por atenuación se elevaban respectivamente a: .73, .74, .79 y .93. Mientras las dos correlaciones temporales significativas de las latencias (uno para los niños y otra para las niñas) que coincidían en el .42 con la corrección sólo aumentaban ligeramente hasta el .47. Los resultados de Gjerde, Block y Block (1985) les llevan a concluir de forma taxativa la clara superioridad de la consistencia de la puntuación de error frente a la de latencias, lo cual les induce también a poner en cuestión a la propia dimensión R-I a la que siguen atribuyendo funciones básicamente de capacidad o maduración intelectual. De todas formas, sin restar ningún mérito al trabajo de Gjerde, Block y Block (1985), en nuestra opinión se deben matizar bastante sus conclusiones. En primer lugar porque tres de las cuatro aplicaciones se han realizado con preescolares (una edad donde ninguna versión del MFF se ha mostrado particularmente útil) y, en segundo lugar, porque a lo largo del estudio se han utilizado tres versiones distintas del MFF de difícil comparación: el MFF-K, que desde 1981 el propio Kagan ya había descartado, el MFF, del cual se hizo una readaptación a 14 ítems sin garantías de mejorar la forma estándar y luego, para la edad de 11 años se construyó una versión propia del MFF con 20 ítems -no confundir con el MFF20 de Cairns y Cammock (1978)-. En resumen, ni hay constancia que las tres versiones sean comparables entre sí, ni es posible comparar el estudio con otros que simplemente han utilizado la forma estándar del MFF de Kagan.

Por lo que respecta a la **fiabilidad del MFF20**, en su trabajo original Cairns y Cammock (1978) lo aplicaron sucesivamente sobre distintas muestras para calcular sus índices psicométricos. Se encontró que la fiabilidad del test utilizando el método de la división por las dos mitades con la corrección de Spearman-Brown era de .89 para los errores y de .91 para las latencias, es decir, bastante más elevada que la del MFF.

También se comprobó que las dos mitades utilizadas, lo que actualmente son los diez primeros y los diez últimos ítems del test, pueden ser utilizadas en la investigación como formas paralelas (Cairns y Cammock, 1978). Por otra parte, el cálculo de la fiabilidad por el sistema de test-retest con un intervalo de cinco semanas confirmó la mejora respecto al test original y ofreció estos resultados: los errores se situaron en el .85 y las latencias en el .77. La correlación negativa entre errores y latencias dependiendo de la edad osciló entre -.62 y -.67 (Cairns y Cammock, 1978, pp. 557-558). Otras características del MFF20 que citan sus autores son, en primer lugar, que no se detectan diferencias por sexos y que las puntuaciones de error y latencias no guardan ninguna relación significativa con el test de inteligencia PMA. En segundo lugar, que la edad ideal de aplicación se sitúa entre los 9 y los 11 años, aunque también resulta útil para niños que tengan entre 7 y 8 años, por contra, para los mayores de 12 años el test pierde mucha capacidad de discriminación. Por último, otro dato importante del estudio de Cairns y Cammock (1978) es que errores y latencias siguen un patrón evolutivo diferente: con la edad disminuyen rápida y significativamente el número de errores, mientras el aumento de las latencias es mucho más pausado.

En el trabajo de Navarro (1987), con aproximadamente 50 sujetos por nivel de edad (1º, 4º, 6º y 8º de Educación General Básica), el coeficiente alfa de consistencia interna del MFF20 fluctuó, en los errores, desde el .50 hasta el .76. Por su parte el de las latencias lo hizo desde el .86 hasta el .93. En el trabajo de Servera (1990) con aproximadamente 25 sujetos por nivel de edad en todos los cursos de Educación General Básica el mismo coeficiente fluctuó entre .60 y .80 para los errores, y entre .93 y .97 para las latencias. En el trabajo más reciente de Buela-Casal, Carretero-Dios, De los Santos-Roig y Bermúdez, M.P (en prensa) con niños también entre 6 y 12 años, pero con muestras de entre 70 a 119 sujetos, los índices de correlación de Spearman-Brown fluctuaron, para los errores, entre .68 y .82 y, para las latencias, entre .92 y .98. En un estudio longitudinal, dentro del mismo trabajo citado, los autores realizaron un seguimiento durante 24 meses de un total de 251 niños de 6 a 9 años en el momento de iniciar el estudio, y ofrecieron coeficientes de

consistencia longitudinal para las puntuaciones conjuntas de errores y latencias (el índice PI de Salkind y Wright, 1977, que explicaremos más adelante) del MFF20 con valores que oscilan entre el .46 y el .61.

En conclusión, parece que el MFF20 presenta mejores índices de fiabilidad que el MFF original, con una puntuación de latencias que ofrece más consistencia interna que la de errores (aunque ambas son elevadas) y con una cierta inversión de los resultados cuando la fiabilidad se mide por procedimientos de test-retest. En cualquier caso llama la atención una cierta disparidad de resultados a través de las edades de los niños y a través de los distintos estudios.

El tema de la **válidez** se ha centrado en la validez convergente y la validez de constructo, si bien hay evidentes interacciones entre ambas. Por otra parte, hay que tener presente que los datos que comentaremos a continuación han sido a menudo criticados por la disparidad de las muestras, la edad de los niños y otros problemas metodológicos en general ya comentados a lo largo de esta monografía.

En lo que respecta a la **válidez del MFF**, los estudios sobre convergencia se basan en la necesidad de que se mantengan las diferencias entre impulsivos y reflexivos en pruebas con parecidas características. Kagan y colaboradores exploraron la validez convergente de las latencias utilizando el ya conocido MFF, junto con el test del Recuerdo Retardado de Dibujos y el test de Emparejamiento Háptico-Visual. Las correlaciones entre los tres tests fueron moderadas y oscilaron entre el .32 y el .52 (Kagan, Pearson y Welch, 1966a). Por su parte, Hall y Russell (1974) también analizaron la validez convergente de las latencias entre el MFF y las matrices progresivas en color de Raven obteniendo una correlación del .54. Yando y Kagan (1970) ofrecen datos de la validez convergente incluyendo a los errores del MFF. Ellos construyeron diez pruebas de emparejamiento de objetos familiares, cada una variaba respecto a las demás en su número de alternativas, con un intervalo que iba desde dos a doce alternativas. Las pruebas fueron administradas a una muestra de niños de 7 años, una cada semana. Los resultados mostraron que a pesar de que el constante aumento de la complejidad de las pruebas provocó un aumento de las latencias y los errores, muchos

niños mantuvieron su rango en la puntuación de errores y latencias respecto a sus compañeros. La correlación media de los errores y las latencias sobre las diez pruebas fue respectivamente de .68 y .73. Denney (1973) encontró que la correlación entre la latencia a la primera respuesta al MFF y la latencia a la primera pregunta en el «*juego de las 20 preguntas*» era del .45.

La correlación por separado de los errores y las latencias con distintas pruebas de solución de problemas ofrece resultados dispares y a veces contradictorios. En general los errores presentan correlaciones más elevadas que las latencias con variables de personalidad, de rendimiento, de conducta motora o simplemente con tests convencionales. Weintraub (1973) con una muestra de 92 niños (normales y en tratamiento clínico) entre 11 y 13 años refiere correlaciones significativas entre estas cinco puntuaciones: errores y latencias del MFF, puntuación de cantidad y de calidad en los laberintos de Porteus y en la latencia de una prueba de demora de la gratificación. Sin embargo, el hecho que todas las puntuaciones correlacionasen con la clase social, tres con la edad, una con el CI y que la muestra estuviera compuesta por niños normales y en tratamiento obliga a matizar mucho los resultados. De hecho, en otros trabajos estos resultados no se han confirmado, por ejemplo, Torner y cols. (1977) trabajando con preescolares no encontró relaciones significativas entre los errores y las latencias de una versión abreviada del MFF, la DALIS (una prueba que consiste en dibujar lo más lentamente una línea sin pararse ni levantar el lápiz) y la WALIS (una prueba similar a la anterior, sólo que esta vez el niño debe caminar lo más lentamente posible), aunque los errores del MFF sí correlacionaron con las pruebas de demora de la gratificación y con las de resistencia a la tentación. En conclusión, a pesar de que los resultados a veces son contradictorios y que se intuye una desigual contribución de las dos puntuaciones del MFF (presentado mejores índices los errores), en general se tiende a aceptar una cierta validez convergente de la R-I muy dependiente de la necesidad que las pruebas presenten incertidumbre de respuesta (Messer, 1976; Palacios, 1982).

Los estudios sobre la **validez de constructo utilizando el MFF** son relativamente escasos y de difícil comparación entre sí. Como ya hemos visto, las conclusiones dependen mucho de la conceptualización que se tenga de la R-I. Para el grupo de Block la dimensión presenta escasa validez puesto que se la examina desde el punto de vista de los estilos cognitivos en su sentido más tradicional. Si se opta por la posición del grupo de Kagan, con una visión más restrictiva de la R-I, más parecida a la que nosotros hemos defendido dentro del campo de los estilos de procesamiento de la información o estrategias de aprendizaje, se han encontrado resultados más favorables.

El primer punto a examinar en cualquier prueba de evaluación de la R-I es el grado de su correlación entre errores y latencias: es crucial que esta correlación sea negativa y elevada para asegurar un buen sistema de clasificación de los sujetos, ya sea por el sistema tradicional que hemos comentado o por otros que comentaremos más adelante. En el caso del MFF se han publicado resultados muy dispares pero hay cierta coincidencia. en las revisiones de Palacios (1982) y Bornas y Servera (1996) en que la correlación entre sus dos puntuaciones está alrededor del $-.50$, tal y como ya apuntaron de forma crítica Egeland y Weinberg (1976) y Ault, Mitchell y Hartman (1976). Con estos niveles, y aplicando además el controvertido sistema de clasificación por cuadrantes, puede ocurrir fácilmente que en dos aplicaciones casi simultáneas del MFF un 25% de sujetos puedan cambiar de cuadrantes o, como se comprobó en el estudio de Egeland y Weinberg, sólo un 50% de impulsivos se mantengan en la misma situación en la segunda aplicación.

Otro punto importante en el estudio de la validez de constructo es observar la relación entre las medidas del MFF y otras pruebas de evaluación que también se dirigen a la evaluación de la «impulsividad», eso sí, ya sea como rasgo de personalidad, social o motora. En este sentido, entre las pruebas de laboratorio que en general hacen referencia a la impulsividad motora, destacan las siguientes: el test de laberintos de Porteus, las tareas de inhibición motora (fundamentalmente la DALS y la WALs ya comentadas), las tareas sociales de toma de decisiones y las

tares de resistencia a la tentación o demora de la gratificación. Por otra parte, entre los tests de personalidad que incluyen la variable impulsividad destacan los siguientes: el EPI y el EPQ de Eysenck (véase Glow y cols., 1983 para una revisión de su relación con el MFF), la «*Impulsivity Scale from the Personality Research Form*» (Davidson y House, 1978), la «*Impulsivity Scale for Children*» (Vacc y Mercurio, 1977) y la «*Auditory Impulsivity Scale for Children*» (Kennedy y Butter, 1978). Entre las escalas comportamentales de evaluación para profesores que valoran una impulsividad social destacan las siguientes: la «*Self-Control Rating Scale*» (Kendall y Wilcox, 1979), la «*School Behavior Checklist*» (Peters y Bernsfeld, 1983), el «*Classroom Behavior Inventory*» (McKinney, 1975) y la «*Conners Teacher Rating Scale*» (Finch, Saylor y Spirito, 1982; Brown y Wyne, 1982). En general las correlaciones del MFF con los tests de personalidad y las escalas para profesores han sido muy poco significativas, tal vez, como apuntan Milich y Kramer (1984), porque la impulsividad cognitiva no tiene porqué coincidir con la impulsividad social o conductual. Que de hecho es la idea que más prevalece hoy en día. Como ya hemos comentado al parecer el mayor ámbito de influencia de la R-I se centra en el rendimiento académico y las tareas escolares. Veamos algunos ejemplos de trabajos que analizan validez de constructo de la dimensión.

Bentler y McClain (1976) realizan un diseño multirasgo-multimétodo complejo e interesante. La muestra es de 68 niños de primer grado y todos ellos son evaluados sobre tres medidas de la impulsividad: MFF, DAL5 y WAL5. Además el diseño incluía tres escalas (una para profesores, otra para los niños y otra para los experimentadores) que evaluaban cuatro dimensiones de la personalidad: extraversión, ansiedad, impulsividad y motivación de logro académico. Al final, los resultados son de difícil interpretación hasta para los propios autores: ninguna de las dimensiones de personalidad de las tres escalas correlacionó con las puntuaciones del MFF. La DAL5, al contrario de la WAL5, correlacionó con los errores y las latencias del MFF. Entre las escalas de personalidad, la DAL5 y/o la WAL5 se dieron algunas correlaciones significativas pero bajas. En su reflexión final los autores se inclinan a pensar que pro-

blemente sus resultados apoyan la concepción inicial de Kagan, en el sentido que la R-I es un fenómeno que actúa sólo bajo determinadas condiciones (por ejemplo, en casos de incertidumbre de respuesta) dentro del paradigma de los procesos de solución de problemas y, por tanto, intentar demostrar su validez de constructo a través de sus implicaciones en el área clínica o de personalidad no es el camino adecuado. Los trabajos de Paulsen y Johnson (1980) y Kendall y Wilcox (1979) presentan conclusiones relativamente en la misma línea.

En uno de los trabajos más citados, Milich y Kramer (1984) afirman que la incapacidad para establecer conclusiones definitivas sobre la validez de constructo de la R-I puede ser atribuible a la falta de control sobre determinadas variables y a la debilidad metodológica. Para intentar subsanar estas deficiencias realizan un estudio con una muestra inicial de 100 niños, de entre los cuales seleccionan 34 por presentar un nivel similar de problemas de aprendizaje y de conducta en el aula. A continuación se controla CI (WISC), edad (6 a 10 años) y sintomatología patológica (entrevista clínica), y se les administran las siguientes pruebas que en total suponen diez medidas de evaluación: MFF, una tarea de resistencia a la tentación (latencias), los laberintos de Porteus, el juego del detective, la DALs y la DALF (una versión contraria a la DALs que consiste en dibujar lo más deprisa posible una línea entre dos paralelas controlando errores y tiempo). Además se pasa un cuestionario de impulsividad de 32 ítems al cual responden: madres, maestros, psicólogos (que habían diagnosticado a los niños inicialmente) y examinadores que están con los niños en situaciones de juego. Los resultados más destacados indican que siete de las diez medidas correlacionan significativamente con la edad y, lo que es aún más preocupante, cinco con el CI (entre ellas los errores del MFF). La matriz que abarca todas las puntuaciones de las diferentes pruebas se compone de 45 correlaciones, de las cuáles 25 son significativas e irían a favor de una cierta consistencia de la validez convergente y de constructo de la R-I. Sin embargo, si se obtienen las correlaciones parciales controlando el CI y la edad únicamente se mantienen como significativas 10 correlaciones.

Por último, Milich y Kramer todavía realizan un nuevo análisis de los resultados siguiendo el principio de agregación de Rushton, que viene a significar que la suma de un conjunto de múltiples medidas de un mismo constructo es un indicador más estable y menos sesgado que el que proporciona cada una de estas medidas por separado. De esta manera, los autores construyen un índice global de impulsividad combinando todas las medidas antes mencionadas y lo correlacionan con las puntuaciones de los cuatro cuestionarios. Los resultados indican que la única correlación significativa se da entre el índice de impulsividad global y el cuestionario de los psicólogos, con un aceptable .62. A modo de resumen final, Milich y Kramer (1984, p. 89) afirman que:

«Después de años de estudio y cientos de artículos, todavía no tenemos una definición válida y aceptable de lo que es la conducta impulsiva. También desconocemos si es posible descubrir un constructo de impulsividad unitario o si por contra debemos concentrarnos sobre diferentes campos de la impulsividad para identificar diferentes aspectos de su comportamiento».

Los autores concluyen que la R-I presenta una baja validez de constructo, de acuerdo con la posición del grupo de Block. Pero evidentemente las conclusiones podrían ser distintas si aceptamos los distintos tipos de impulsividad y justificamos así la falta de correlación del MFF con medidas motoras y de comportamiento social. Tal vez el dato más preocupante de los recogidos por estos autores en la significativa relación entre la R-I y el CI, pero anteriormente ya hemos matizado suficientemente esta posibilidad que, en cualquier caso, sigue siendo un foco de polémica.

Por otra parte, la relación entre el MFF (o el MFF20 puesto que aquí no hay gran diferencia entre las dos versiones) y medidas de rendimiento académico, atendiendo a la visión que enlaza la R-I con estilos de procesamiento de información o estrategias de aprendizaje más o menos adaptativas, es el dato más consistente y significativo de todas las líneas de investigación. En efecto, desde que en 1971 Kogan ya destacase que comparado con otros nueve estilos cognitivos, incluida la dependencia-independencia de campo, la R-I era el que guardaba más re-

lación con el ámbito educativo, la lista de estudios que han podido constatar que los niños reflexivos tienen un mejor rendimiento que los niños impulsivos es muy amplia (Keogh y Donlon, 1972; Finch y cols., 1974; Barret, 1977; Grinberg, 1980; Margolis, 1982; Weithorn, Kagen y Marcus, 1984; Shorr y Dale, 1984; Brown y Wynne, 1984, etc.).

La **validez de constructo del MFF20** es más sólida al menos en lo que respecta a la correlación entre errores y latencias. En el trabajo original de Cairns y Cammock (1978), aunque las muestras para cada nivel de edad fluctuaban de entre 37 a 95 sujetos, ya se recogieron correlaciones de siempre cercanas o superiores al $-.60$. En el trabajo de Navarro (1987), si exceptuamos la muestra de 8º de Enseñanza General Básica, las correlaciones fueron iguales o superiores a $-.68$. En el trabajo de Servera (1990), también dejando de lado los niños de 8º, fueron iguales o superiores a $-.60$. Los datos más recientes de Buela-Casal, Carretero-Dios, De los Santos-Roig (en prensa) con muestras superiores a 50 niños por grupo de edad (6, 7, 8 y 9 años) siguen en la misma línea: los coeficientes fluctúan entre $-.52$ y $-.75$, siendo el valor alrededor del $-.60$ el más habitual. En definitiva, tal vez estamos hablando de una mejora entre 10 y 15 puntos en relación a los coeficientes de correlación del MFF.

En lo que respecta a la relación del MFF20 con otras medidas afines al constructo impulsividad los datos, en general, no difieren muchos de los hallados con el MFF. Esto es, muy poca relación entre la R-I y los cuestionarios o escalas de personalidad y comportamiento y mucha más relación con las medidas de rendimiento académico (Servera, 1990; 1992). En un interesante trabajo citado en Buela-Casal y cols. (2001a) se presentan los datos de una muestra de 74 sujetos entre 11 y 12 años evaluados a través del Cuestionario de Personalidad para Niños de Porter y Cattell y el MFF20. Los resultados muestran que prácticamente no existe ninguna relación entre la R-I y las 14 dimensiones primarias de personalidad que mide el cuestionario, a excepción hecha de la de dominante-sumiso y la más previsible de inteligencia alta-baja.

En definitiva, a modo de conclusión, cabe destacar en lo que respecta a los aspectos de fiabilidad que el MFF20, aún siendo mucho menos utilizado por la comunidad científica, presenta mejores índices que el MFF, si

bien bastante variables según se estudien por procedimientos de consistencia interna o por test-retest. También con algunos problemas metodológicos, los datos apuntan a una cierta consistencia longitudinal de la dimensión. Por lo que se refiere a los aspectos de validez, también el MFF20 ha conseguido mejorar la correlación errores-latencias, imprescindible para una buena clasificación de los sujetos. La validez convergente y de constructo del MFF y del MFF20 resulta más problemática: ciertamente no hay otros sistemas de medida que eliciten tan claramente la R-I (lo cual es un problema relativamente importante) y, además, no se ha encontrado relación entre este tipo de impulsividad cognitiva y otras de carácter comportamental, de personalidad o motora. Para los autores que defienden una concepción pura, en estilo cognitivo, de la R-I estos datos invalidan la dimensión, pero para los que defienden una concepción más restrictiva, centrada en aspectos de aprendizaje y funcionamiento cognitivo, no resultan tan negativos. Después de todo es sobre medidas de rendimiento donde estos estilos de procesar la información deben mostrar su adecuación, sin necesidad de depender de factores de capacidad o inteligencia.

2.3.2 Sistemas de clasificación en los MFF: el modelo integrado de Salkind y Wright

A lo largo de los anteriores apartados hemos ido exponiendo implícita o explícitamente críticas conceptuales y metodológicas a la dimensión R-I. Mientras las conceptuales, al menos a nuestro juicio, tienen mejor defensa, las metodológicas son más conflictivas. Estas críticas metodológicas se centran, por un lado, en el test MFF y, por otro lado, en el sistema de clasificación basado en la doble división por las medianas de errores y latencias. En realidad hay que aceptar que son muchos los autores que coinciden en rechazar tanto al MFF por sus limitaciones psicométricas como al sistema de clasificación tradicional por su ineficacia para validar la dimensión R-I (Ault, Mitchell y Hartman, 1976; Buela y cols., 2001a, Margolis, Peterson y Leonard, 1979; Grinberg, 1980; Victor, Halverson y Montague, 1985; Salkind y Wright, 1977; Gjerde, Block y Block, 1985; Palacios, 1982; Navarro, 1987; Quiroga y Forteza; 1988; Servera 1990, 1992).

En cierto modo, aunque con algunas limitaciones que justificaremos más adelante, al abordar los objetivos de la parte experimental del presente trabajo, el MFF20 vino a subsanar parte de los problemas. Sin embargo, lo que hace referencia al sistema de clasificación quedó en el aire. Como ya comentamos, el sistema de clasificación por cuadrantes está afectado de una gran arbitrariedad, debido esencialmente a que se rige por un estadístico meramente descriptivo (la mediana) para intentar realizar una clasificación de carácter cualitativo. La clasificación por el sistema tradicional también obliga a realizar un acto con pocos precedentes en el campo de las dimensiones psicológicas, y concretamente en el de los estilos cognitivos: casi un tercio de las muestras son descartadas al no poder asignarles una puntuación en impulsividad. Pero además estos sujetos descartados, lentos-inexactos y rápidos-exactos, a menudo presentan más diferencias e implicaciones que los propios impulsivos y reflexivos. En definitiva parece evidente que la inmensa mayoría de problemas que afectan al sistema de clasificación tradicional derivan de una sola cuestión: la imposibilidad de trabajar con puntuaciones continuas y de separar los componentes estilístico y de capacidad que se conjugan en la realización de pruebas como el MFF. La solución a este dilema es el llamado **«sistema de clasificación integrado»** de Salkind y Wright.

Fue en 1977 cuando Neil Salkind y John Wright publican su trabajo, ya clásico, en el que exponen el sistema de clasificación integrado, que supuso además un gran impulso para la R-I. En la parte introductoria ellos reconocen que las puntuaciones de error y latencias sobre el MFF presentan un sesgo hacia los errores, tal y como señaló el grupo de Block. Sin embargo, en su opinión las latencias, a pesar de tener una consistencia temporal más baja y guardar menos relaciones con otras medidas, son un tipo de medida que ciertamente podría estar más relacionada con las expectativas del individuo, sus disposiciones o su estrategia de trabajo en un contexto relativamente independiente de los factores de capacidad. Según Salkind y Wright (1977) ni la utilización por separado de las puntuaciones de errores y latencias, ni la *«suma»* conjunta al estilo del sistema tradicional responden a la definición y a la idea conceptual de la R-I. Por tanto, la postura del grupo de Block es

acertada al rechazar dicho sistema de clasificación pero, en cambio, es errónea al referirse a las latencias como un índice trivial. Tal y como ya se puso en evidencia con trabajos iniciales sobre la noción de «*estilo cognitivo*» en última instancia las puntuaciones de error y de latencias sobre cualquier tarea dependen de las características de la propia tarea y del modo de trabajar del sujeto. Otros trabajos han podido constatar esta evidencia, entre ellos están, por ejemplo, los de Rollins y Genser (1977), Zelniker y Jeffrey (1976, 1979) y Smith y Kemler-Nelson (1988).

La solución Salkind y Wright para disponer de una puntuación continua en estilo y otra totalmente independiente en capacidad sobre los tests MFF tiene una fuerte base matemática. Los propios autores resumen conceptualmente el proceso aplicado como...

«...una rotación de los cuatro cuadrantes de la distribución original (es decir, surgida del sistema tradicional) de forma que el índice de impulsividad se define como una dimensión de las diferencias individuales clasificadas desde las actuaciones rápidas-exactas a las lentas-inexactas. El índice de eficacia se define como una dimensión conceptualmente ortogonal a la impulsividad, a lo largo del cual las diferencias individuales son clasificadas desde las actuaciones lentas-inexactas a las rápidas-exactas» (Salkind y Wright, 1977, p. 381).

En la práctica esta rotación supone definir dos puntuaciones continuas independientes entre sí: una puntuación en estilo o impulsividad (PI) y otra puntuación en capacidad o eficacia (PE), tal y como explica la siguiente tabla:

Tabla 1. El sistema de clasificación integrado de Salkind y Wright (1977)

Puntuación en Impulsividad (PI):	PI = Zerrores - Zlatencias
Puntuaciones negativas (-) Tendencia al polo reflexividad	Puntuaciones positivas (+) Tendencia al polo impulsividad
Puntuación en Eficacia (PE):	PE = Zerrores + Zlatencias
Puntuaciones negativas (-) Tendencia al polo eficacia	Puntuaciones positivas (+) Tendencia al polo ineficacia

Para llegar a calcular la PI y la PE es necesario evaluar una muestra de niños, por ejemplo los pertenecientes a una aula escolar, y a continuación calcular la media y desviación estándar de sus puntuaciones de errores y latencias. A continuación las puntuaciones directas de cada niño se transforman en puntuaciones Z o típicas (aquellas cuya distribución resultante siempre tiene una media 0 y una desviación estándar 1) siguiendo el procedimiento acostumbrado, esto es, a cada puntuación directa se le resta la media de la distribución muestral y el resultado se divide por la desviación estándar. Cuando ya tenemos todas las puntuaciones directas de errores y latencias transformadas a puntuaciones Z simplemente basta con restarlas para conseguir la PI y sumarlas para conseguir la PE. Si se realiza una comparación entre el sistema tradicional y el integrado se puede comprobar que:

1. Prácticamente todos los sujetos impulsivos del sistema tradicional siguen siendo impulsivos (es decir, con PI positivas) en el sistema integrado. Cuanto más alejada del cero es la puntuación positiva mayor es el grado de impulsividad que se infiere.
2. Prácticamente todos los sujetos reflexivos del sistema tradicional siguen siendo reflexivos (es decir, con PI negativas) en el sistema integrado. Cuanto más alejada del cero es la puntuación negativa mayor es el grado de reflexividad que se infiere.
3. En el sistema integrado los sujetos lentos-inexactos y rápidos-exactos del sistema tradicional serán «*reflexivos*» si su puntuación estándar de latencias es mayor que la de errores y serán «*impulsivos*» si es al revés. Aunque normalmente ocupan lugares intermedios en lo que es el «*continuum*» de la PI.

Por su parte, la PE es, por definición, ortogonal a la PI. Es decir, su coeficiente de correlación está alrededor del cero. Pero igualmente mantiene la lógica del sistema tradicional: como sería de esperar prácticamente todos los sujetos «*rápidos-exactos*» de este sistema son clasificados como «*eficaces*» (PE negativa) y todos los sujetos «*lentos-inexactos*» del sistema tradicional ahora presentan PE positivas. Los sujetos

impulsivos y reflexivos del sistema tradicional normalmente ocupan posiciones intermedias en el «*continuum*» de la PE, pero saber a primera vista si serán clasificados como «eficaces» o «ineficaces» es algo más complicado. Un modo de averiguarlo es prescindir del signo de las puntuaciones estándar y fijarse solamente en su tamaño en términos de valor absoluto: los sujetos reflexivos que presentan una puntuación estándar de latencias menor que la de errores entran a forma parte del polo de eficacia y, si es al revés, aparecen como ineficaces. Por su parte, los sujetos impulsivos que presentan una puntuación estándar de latencias mayor que la de errores (siempre en términos de valor absoluto) aparecen como eficaces, mientras si es al revés aparecen como ineficaces.

En definitiva, al tratarse de un sistema matemático, presenta sus limitaciones inherentes a los cálculos y a veces, para realizar interpretaciones correctas, hay que ser precavidos, sin que hasta el momento se hayan dado normas específicas de cómo hacerse. Por ejemplo, se supone que la validación del sistema integrado está íntimamente ligada a la propia validación de la R-I. En principio el sistema hipotetiza que los índices PI y PE presentarán correlatos con variables de naturaleza diferente: la PE mantendrá una mayor relación con pruebas o factores de maduración y capacidad cognitiva, mientras la PI deberá presentar más implicaciones con todo lo que sean medidas o factores estratégicos o de estilo de afrontamiento de las tareas.

En apoyo del modelo de Salkind y Wright (1977) podemos destacar los trabajos de Kogan (1976), Grinberg (1980) y Quiroga y Forteza (1988) que han encontrado que la PE mantiene correlaciones significativas con la variable edad, de modo que la eficacia se relaciona con la maduración física e intelectual del niño. En cambio la PI presenta más estabilidad al aumento de edad. Además, Quiroga y Forteza (1988) con una muestra de 48 niños y niñas entre 8 y 10 años encontraron que mientras el CI (medido con cinco escalas del WISC) mantenía correlaciones bajas y similares con los errores y las latencias (respectivamente de $-.12$ y $-.18$), éstas diferían en signo y calidad respecto a las PI y PE, que eran respectivamente de $.05$ y $-.30$. En la misma línea, Miyakawa

(1981) con una muestra japonesa de 144 sujetos entre 9 y 10 años también encontró que la PE mantuvo correlaciones (la mayoría significativas) más elevadas que la PI tanto con la edad como con las escalas de inteligencia del «*Kyoto University NX Group Intelligence test*».

Aunque de modo más discutible a favor del sistema de Salkind y Wright (1977) también está el trabajo de Servera (1990). En este trabajo sobre un cuestionario para maestros que evalúa variables de rendimiento y de conducta de los niños se compara el sistema de clasificación tradicional con un sistema de cuadrantes alternativo surgido de combinar la PI y la PE. En principio los resultados indican que los dos sistemas son capaces de diferenciar el rendimiento académico en la dirección esperada, es decir, los reflexivos son el grupo con mejor rendimiento y los impulsivos el de peor rendimiento, sin embargo, con el sistema de cuadrantes PI/PE las diferencias alcanzan un mayor nivel de significación y además se extienden a variables de comportamiento «*estratégico*». Únicamente los reflexivos del sistema PI/PE se diferencian claramente de los impulsivos en la «*impulsividad*», la «*concentración*» y la «*mayor persistencia ante tareas difíciles*» según la valoración de los maestros. Por otra parte, en el mismo trabajo a través de los coeficientes de determinación se analiza la capacidad de predicción de los errores, las latencias la PI y la PE sobre las variables de rendimiento académico. Los coeficientes son relativamente bajos pero muestran a la PI como la principal variable (predice alrededor del 10% de la variabilidad de las puntuaciones en matemáticas, lengua y en rendimiento global del niño), seguida de los errores (alrededor del 5%), mientras las latencias y la PE presentan valores insignificantes.

Al igual que en el trabajo de Servera (1990) se echa en falta una mayor relación de la PE con variables de rendimiento, en otros trabajos también hay que destacar datos poco favorables al modelo de clasificación integrado. Looper y Hallahan (1980) encontraron que la PI, los errores y las latencias mantenían relaciones significativas con el rendimiento en lectura, pero ése no era el caso de la PE. El propio Salkind (1975) constató que las correlaciones de la PI con el PMA de Thurstone, a pesar de ser algo más bajas que las de la PE, eran estadís-

ticamente significativas. Miyakawa (1981), en el trabajo ya citado, descubrió que con una muestra de niños de primer grado sus resultados iniciales con niños mayores se invertían y era la PI la que mantenía las correlaciones más elevadas con la edad y con las pruebas de inteligencia. Navarro (1987) no encontró relaciones importantes entre la PE y la puntuación de las Matrices Progresivas de Raven. Por último, Servera (1992) con una muestra de 81 niños y niñas entre 10 y 11 años encontró resultados algo más favorables al sistema integrado: por un lado el Raven correlacionó significativamente con los errores (-.38) pero no con las latencias (.09) y, por otro lado, mostró una correlación algo superior con la PE (.32) que con la PI (-.26), aunque ambas estadísticamente significativas.

El trabajo de Buela-Casal, De los Santos-Roig y Carretero-Dios (en prensa) demuestra como los supuestos básicos del modelo de Salkind y Wright se cumplen aplicando la adaptación española baremada del MFF20, pero no aportan datos respecto a la distinta significación de la PI y la PE. En otro trabajo de los mismos autores (Buela-Casal, Carretero-Dios y De los Santos, 2000) se analiza en un estudio longitudinal la relación de la R-I con el rendimiento académico. El objetivo es determinar la validez criterial de la dimensión y ello en principio se consigue puesto que los reflexivos presentan un significativo mejor rendimiento que los impulsivos y, afinando más, al incorporar la PE al estudio se observa que el orden de mejor a peor rendimiento es el siguiente: reflexivos-exactos, reflexivos-ineficientes, impulsivos-eficientes e impulsivos-ineficientes. De modo que prevalece la PI sobre la PE. Aunque no aparecen referencias directas a la correlación de ambas puntuaciones sobre las distintas medidas de rendimiento es fácil suponer, como en los restantes estudios citados, que con toda probabilidad la PI presentaría siempre correlaciones más elevadas que la PE. Esto es un dato que para muchos autores no favorece ni al modelo integrado ni a la propia consistencia de la R-I: ¿cómo puede ser que si la PE es un índice de capacidad presente correlaciones más bajas que la PI, una puntuación «estilística», en pruebas de rendimiento académico o tareas de resolución de problemas en general?

No existe una respuesta clara para esta pregunta, salvo si se rechaza el modelo integrado. Pero ello no parece lo más adecuado, porque como comentan Buela-Casal y cols. (2000), realmente se demuestra una elevada validez criterial de la R-I aplicada al rendimiento académico, que de alguna manera debe explicarse. Tal vez, y aún a riesgo de inferir bastante más de lo que los datos nos permiten, los tests MFF son capaces de establecer diferencias mucho mayores en lo que a estilo se refiere que en lo que es estrictamente la capacidad de los niños. De hecho, como sabemos todos los que hemos evaluado a miles de niños con esta prueba, salvo si hay un problema grave de deterioro intelectual, todos los niños acaban por encontrar la solución (aunque sea por suerte), y en cambio sólo les hemos tomado en consideración su latencia a la primera respuesta. Si realmente quisieramos evaluar la capacidad intelectual del niño con este tipo de pruebas buscaríamos ítems más difíciles (en casi todas las versiones se eliminan los ítems más fáciles y más difíciles cuando se construyen las pruebas), y mediríamos el tiempo total o las latencias a cada respuesta para ponderar mucho más este índice. O, como propone el modelo probabilístico de Solís-Cámara (1987), antes de efectuar ninguna clasificación intentaríamos dejar fuera de las muestras a aquéllos niños que han mostrado un comportamiento azaroso sobre el MFF y sólo trabajaríamos con aquéllos que realmente han adoptado una estrategia de trabajo, sea impulsiva o reflexiva.

En cualquier caso, y a pesar de los problemas comentados, el modelo de Salkind y Wright presenta ventajas muy a tener en cuenta: permite separar el componente de eficacia del de estilo contemplando la mutua influencia de las puntuaciones de error y latencias, permite asignar a todos los sujetos de una muestra una puntuación en impulsividad y otra en eficacia sin necesidad de eliminar ninguno, permite trabajar no sólo con el análisis de variancia sino también con los modelos de regresión al ser la PI y la PE variables continuas y, por último, permite (o debería permitir) facilitar la posterior normalización de las puntuaciones.

Por último, pero no menos importante porque enlaza directamente con los objetivos de nuestro trabajo experimental, hay un tema que ha pasado por alto pero que a nuestro juicio puede tener gran relevancia.

¿Cómo determinar quien es realmente impulsivo o reflexivo en una puntuación continua tipificada?. Es evidente, como comentamos ya, que en el punto intermedio, alrededor del valor 0 de la PI, se concentran sujetos rápidos-exactos y lentos-inexactos o simplemente aquéllos a los que no resulta nada claro asignarles una tendencia estilística. Servera (1990) o Buela-Casal y cols. (2000) han optado por dos vías: una es tomar ± 1 desviación estándar sobre la PI para clasificar a sujetos claramente impulsivos de los claramente reflexivos, con independencia de la PE. Esta vía supone un problema importante: por ejemplo, ¿comparten las mismas características estilísticas dos sujetos con una PI de +1,5 (es decir, claramente «impulsivos») si además uno presenta una PE de -1,5 (claramente eficaz) y otro de +0,5 (prácticamente sobre la media exacta de eficacia/ineficacia)?. Nosotros, como justificaremos más adelante, pensamos que no, y que ello debe tenerse presente a la hora de establecer las clasificaciones cualitativas (no las meramente cuantitativas) sobre la R-I.

La segunda vía es combinar las puntuaciones PI y PE, estableciéndose así un nuevo sistema de cuadrantes pasado por el filtro de la tipificación: ahora, en función de si están por encima o por debajo del valor 0, tenemos sujetos reflexivos-eficaces, reflexivos-ineficaces, impulsivos-eficaces e impulsivos-ineficaces. Pero no hemos avanzado mucho: en el mismo cuadrante (por ejemplo, el impulsivo-ineficaz) pueden estar un sujeto con puntuaciones de PI +0,1 y de PE +0,1 con otro de puntuaciones mucho más extremas de PI +3 y de PE +1, lo cual, como en el ejemplo anterior, tampoco parece en absoluto justificado. La solución a este problema la propondremos más adelante en nuestro estudio experimental.

2.4 El aprendizaje y el comportamiento del «niño impulsivo»

Hasta ahora hemos revisado la conceptualización, la forma de evaluación y los sistemas de clasificación basados en la dimensión Reflexividad-Impulsividad. El objetivo del presente apartado es bajar al ruedo escolar, a lo que es el ámbito normal donde se mueven los niños y definir las características que padres, educadores y psicólogos pue-

den apreciar en un «niño impulsivo», o hablando más en propiedad, aunque nos tomemos la licencia de utilizar ambas expresiones de forma sinónima, un niño de estilo tendente a la impulsividad.

Ha quedado bastante claro a estas alturas que, tanto desde el punto de vista conceptual como por los datos empíricos, las características distintas del niño impulsivo son mucho más evidentes en el ámbito del aprendizaje que en el ámbito comportamental. Hasta el punto que probablemente el mejor dato de validez criterial de la R-I es cuando se analizan sus relaciones con las notas escolares o tareas de rendimiento. Existen muchos datos con el propio MFF que sirven para destacar que los reflexivos acostumbran a demostrar un mejor rendimiento que los impulsivos tanto sobre materias de matemáticas, geometría y ciencias en general, como muy especialmente en lectura y comprensión lectora. El lector interesado puede hallar más información y múltiples referencias sobre esta constante en las revisiones de Bornas y Servera (1996), Palacios (1982), Navarro (1987), Servera (1992). Veamos ahora algunos estudios específicos.

Servera (1990) con una muestra de 197 alumnos que se repartía de forma casi equitativa entre todos los cursos de lo que era la Enseñanza General Básica pudo comparar la influencia de la R-I, la dependencia-independencia de campo (DIC) y el ITPA (el Test de Aptitudes Psicolingüísticas de Illinois) sobre el rendimiento escolar evaluado por los profesores. La tabla 2 muestra los resultados de las correlaciones halladas.

Tabla 2. Resultados de las correlaciones de las medidas de la R-I (MFF20), la DIC y el ITPA sobre el rendimiento académico

Rendimiento académico	Errores MFF20	Latencias MFF20	PI (p. en impulsividad)	PE (p. en ineficacia)	DIC*	ITPA**
Global	-,22	,07	-,28	-,22	,15	,43
Matemáticas	-,20	,15	-,27	-,15	,14	,32
Lenguaje	-,25	,12	-,31	-,19	,08	,34

* Se utilizó el CEFT para los cursos de 1º a 4º y el EFT para los cursos de 5º a 8º.

** Sólo se aplicó a los cursos de 1º a 3º.

Los resultados muestran correlaciones siempre en la dirección esperada: el mejor rendimiento se asocia con la «*reflexividad*», la «*independencia de campo*» y las «*aptitudes para el aprendizaje*» del ITPA. Como era de esperar ésta fue la prueba que presentó mayor relación con el rendimiento académico, pero a continuación hay que destacar que las medidas de la R-I tienen más influencia que las de la DIC, hecho por otra parte ya constatado por Kogan (1971). Dentro de la R-I, los errores ofrecen correlaciones con el rendimiento mucho más elevadas que las latencias, pero la PI las llega a superar, relegando incluso a la PE a un segundo plano. De este modo el estilo impulsivo demuestra tener algún tipo de incidencia directa sobre el rendimiento académico. Algunos pueden pensar que correlaciones alrededor del $-.30$ son excesivamente bajas. Pero si pensamos en términos de coeficientes de determinación (es decir, al elevar al cuadrado las correlaciones se obtiene matemáticamente el porcentaje de variabilidad que una variable es capaz de explicar) nos encontramos con una realidad diferente. En nuestro caso, como en el de otros muchos estudios, resulta que más o menos el 10% de las puntuaciones en rendimiento de los niños son explicadas por la PI, una puntuación que se ha extraído, no lo olvidemos, de una prueba de discriminación perceptiva que se aplica en unos 10 minutos. Si tenemos en cuenta que las complejas baterías de inteligencia se dan por satisfechas (como ha pasado aquí con el ITPA) si pueden explicar entre el 20-25% del rendimiento académico del niño desde luego no cabe pensar en las correlaciones de la PI como algo baladí.

En otro estudio similar pero mucho más específico Servera (1992) comparó las puntuaciones del MFF20 en una muestra de 81 niños de entre 10 y 11 años con los resultados obtenidos por éstos en medidas objetivas de rendimiento académico extraídas de su curriculum escolar (no eran, pues, simples valoraciones de sus profesores). Aunque los coeficientes de correlación volvieron a ser moderados, como no puede ser de otra manera, fueron suficientes como para que se descubrieran diferencias estadísticamente significativas entre impulsivos y reflexivos. Lo importante del dato es que estas diferencias fueron independientes del sistema de clasificación utilizado, basado en la PI o basado en los cuadrantes y, lo que aún resulta

más destacable, no fue necesario comparar grupos extremos para detectar ya diferencias significativas. La tabla 3 muestra estos resultados.

Tabla 3. Diferencias entre niños impulsivos y reflexivos en medidas objetivas de rendimiento académico (R.A.)

	Clasificación por el sistema integrado (PI)				Clasificación por el sistema de cuadrantes			
	Impulsivos (n=43)		Reflexivos (n=38)		Impulsivos (n=30)		Reflexivos (n=26)	
R.A.*	Media	DE	Media	DE	Media	DE	Media	DE
Compr. lectora	2,86	1,92	4,39	2,51	2,70	1,49	4,73	2,44
Matemáticas	4,60	2,39	5,87	2,47	4,73	2,41	6,46	2,31
Expres. escrita	2,95	1,56	4,47	1,81	2,80	1,40	4,88	1,80
total**	11,02	5,34	15,97	6,35	10,70	4,40	17,31	5,86

Nota: Todas las diferencias son estadísticamente significativas con una $p < .01$, excepto en la puntuación de matemáticas para la clasificación por PI en la que la $p < .05$.

* La puntuación en comprensión lectora, matemáticas y expresión escrita varía entre 0 y 10.

** La escala total incluye un total de nueve medidas de evaluación y varía de 0 a 33 puntos.

Las conclusiones más destacadas de este trabajo, a parte de la superioridad en rendimiento académico de los niños reflexivos sobre los impulsivos, son, por una parte, que al no haber diferencias muy importantes entre los sistemas de clasificación, parece mucho más interesante optar por el de la PI que no nos obliga a prescindir de ningún niño (los sujetos rápidos-exactos y lentos-inexactos quedan fuera en la clasificación por cuadrantes). Por otra parte, como también se analizó en el trabajo de Servera (1992), clasificar a los sujetos a partir de 1 desviación estándar por arriba y por debajo, es decir, seleccionar a impulsivos y reflexivos más extremos, garantiza que las diferencias que ahora ya son significativas se incrementen todavía bastante más.

En uno de los estudios más completos e interesantes hasta la fecha Buela-Casal, Carretero-Dios y De los Santos-Roig (2000) realizan un análisis longitudinal de las diferencias en rendimiento académico entre reflexivos e impulsivos evaluados con el MFF20. El rendimiento se mide a través de las notas oficiales de los niños en seis momentos tempora-

les: pasados 1, 3, 9, 13, 15 y 21 meses de la evaluación con el MFF20. A los 24 meses se volvió a aplicar el MFF20. Los resultados que más nos interesan aquí son bastante coincidentes con los ofrecidos de la investigaciones propias pero, en general, con coeficientes de correlación más elevados: así, por ejemplo, los errores correlacionan $-.43$ con la media total de rendimiento y las latencias, de modo relativamente sorprendente, llegan hasta el $.25$. La PI se queda en el valor habitual del $-.33$, explicando un 11% del rendimiento académico global. Por supuesto las diferencias entre reflexivos e impulsivos siguen la dirección esperada, demostrándose en este caso que se mantienen en el tiempo prácticamente en todas las materias, con algunas excepciones especialmente en educación artística.

Es necesario remarcar que aunque las diferencias sean estadísticamente significativas en todos los estudios no se trata de diferencias «*astronómicas*», puesto que en realidad éstas sólo cabe esperarlas entre niños «*muy inteligentes*» y niños con «*déficits cognitivos*» más o menos acusados. Pero reflexivos e impulsivos no difieren en este tipo de medidas, por más que se muestre la ligera superioridad de los reflexivos. No hay que confundir, pues, un niño lento-inexacto con niño impulsivo (aunque sus problemas académicos pueden ser similares), puesto que sólo en el primero cabe destacar problemas de desarrollo cognitivo más o menos acusados que, además, se generalizarán a otros ámbitos. ¿Qué cabe esperar, pues, del estilo de aprendizaje de un niño impulsivos?. Para contestar a esta pregunta hay que diferenciar primero lo que sería un niño impulsivo «*puro*», es decir, sin ninguna otra problemática, del que no lo es. Es relativamente fácil encontrar estilos impulsivos en niños con dificultades de aprendizaje graves (dislexias, discálculias), con ligero retraso del desarrollo, con hiperactividad, con situaciones sociofamiliares desfavorecidas, etc. Pero en estos casos la impulsividad estilística sólo es un problema más, no necesariamente el principal. En cambio sí lo es en el niño impulsivo «*puro*». Este niño únicamente presenta problemas visibles sobre el rendimiento académico, pero no es en absoluto conflictivo en casa o la escuela e incluso puede pasar bastante desapercibido. En las tareas escolares comete errores «*ontos*», se precipita a la

hora de responder, parece como si su acción vaya por delante de su planificación, no es sistemático a la hora de recopilar toda la información que le sería necesaria para solucionar los problemas, etc.. En resumen, como ya destacamos en su momento, estamos ante un niño con problemas evidentes en todo el proceso de evaluar y tomar decisiones especialmente en situaciones con incertidumbre de respuesta, y las escolares lo son en muchos casos.

En principio este problema general de estilo es suficiente para dificultarle su capacidad de aprendizaje, no tiene porqué, pues, ir asociado a dislexias, discalculias, déficits de atención, memorísticos, de lenguaje u otras dificultades más específicas; la impulsividad sería, de algún modo, un tipo específico de dificultad de aprendizaje. Por otra parte, y como ya hemos ido señalando a lo largo del trabajo, no cabe esperar que el niño impulsivo se diferencie comportamental de otros, o que su impulsividad se halle ligada a algún trastorno del comportamiento. Sin embargo, aquí enlazamos con una confusión no del todo aclarada: muchos entienden que impulsividad es sinónimo o acompañante de los trastornos por déficit de atención con hiperactividad (TDAH), pero en realidad no es exactamente así.

El TDAH es una alteración neuropsicológica grave que se caracteriza por síntomas alterados en el área atencional y motora (impulsividad y sobreactividad). Aunque los MFF se han incorporado muchas veces a la evaluación de niños sospechosos de padecer hiperactividad bajo la suposición de que la medida de estilo ayudaría al diagnóstico, no han tenido mucho éxito. En resumen, y aunque es un tema polémico, la cuestión es que la impulsividad fundamental de los niños con TDAH es más bien motora o manifiesta, al hilo de los distintos tipos de impulsividad que definimos al principio del presente trabajo. De hecho, en general, la mayoría de niños con TDAH sobre los MFF se muestran lentos-inexactos y no precisamente impulsivos (Firestone y Martin, 1979; Sandberg, 1978, Sergeant, 1981 -citado en Taylor, 1991a-; Swanson, 1984). Ello quiere decir que su signo más distintivo es el cometer múltiples errores, pero sin necesidad de ir rápidos. Al perder muy deprisa la capacidad de mantener la atención y al tener todavía muchos menos recursos de afronta-

miento de tareas que los niños de estilo impulsivo, los niños con TDAH simplemente *«están fuera de la tarea»*.

A pesar de las múltiples controversias, y las limitaciones metodológicas, nos podríamos aventurar a conceptualizar la relación entre estilo impulsivo y TDAH del siguiente modo. En primer lugar, cabe esperar que la inmensa mayoría de niños hiperactivos comentan muchos más errores que los normales sobre los MFF. En segundo lugar, las diferencias con respecto a la puntuación de latencias no están tan claras, en general encontraremos muchos hiperactivos con el patrón *«dento-ine-xacto»*, pero cabe la posibilidad de que otros se muestren *«impulsivos»* (es el caso del trabajo de Furhman y Kendall, 1986 que dan más peso a la PI que a la PE). Por último, y por encima de la anterior consideración, se puede afirmar con rotundidad que la probabilidad de encontrar niños hiperactivos que se muestren impulsivos sobre el MFF es mucho mayor que la de encontrar niños impulsivos que cumplan con los criterios diagnósticos de la hiperactividad. En definitiva, el MFF sólo puede ser una prueba complementaria para la evaluación de niños hiperactivos.

Para finalizar, debemos reiterar que la investigación no ha encontrado más relaciones directas entre impulsividad cognitiva y otras variables del comportamiento. Niños desinhibidos, con poco autocontrol, con conductas disruptivas en el aula pueden ser o no ser de estilo impulsivo, pero nada garantiza una relación ni causal ni siquiera correlacional. En cambio, como hemos visto a lo largo de toda esta parte, se puede detectar a un niño de estilo impulsivo, sin ningún problema comportamental o social destacable, sólo con dificultades, en algunos casos graves, en el ámbito del aprendizaje y el rendimiento académico. Otra cosa es que los problemas subyacentes a esta impulsividad puedan desgranarse en mecanismos más básicos, como la falta de concentración y atención, sin necesidad de llegar a constituir un trastorno tan importante como el TDAH, pero manteniendo evidentes vínculos directos. Para ello no será necesario, al igual que hemos hecho con la impulsividad, analizar la naturaleza multicomponencial de la atención.

3. LA NATURALEZA MULTICOMPONENTE DE LA ATENCIÓN

La psicología del procesamiento de la información, desarrollada y potenciada a partir de los años sesenta, recuperó una de las líneas más clásicas de la psicología: el estudio de la actividad mental. En palabras sencillas pero elocuentes, dicha actividad se sirve de mecanismos, procesos y operaciones para ayudar al organismo a adaptarse al ambiente (García-Sevilla, 1997). Así, inteligencia, percepción, representación, lenguaje, aprendizaje, memoria y, por supuesto, atención forman parte de esta actividad. Sin embargo, como también recalca la autora citada, en los manuales de Psicología General o Básica realmente la atención no parecía estar a la altura de los demás conceptos a tenor del reducido espacio que se le dedicaba. Ello nunca fue óbice para que psicólogos, educadores y personal del ámbito clínico o escolar siempre tuvieran muy presente el término problema de atención para referirse a multitud de conductas, déficits, disfunciones o alteraciones no siempre equiparables, ni de la misma naturaleza.

El presente capítulo servirá para analizar el concepto de atención, y muy especialmente el de atención sostenida, un mecanismo ampliamente vinculado a problemas de aprendizaje y algunos trastornos de conducta, como la hiperactividad. Comparte, en general, las mismas dificultades de conceptualización y medida que la impulsividad cognitiva y, por tanto, forma parte de nuestro interés el desarrollar sistemas de evaluación que no sólo tengan una validez intrínseca, sino que tengan impacto en lo que es la consideración de la capacidad de aprendizaje y de comportamiento del niño en la escuela.

3.1 Conductas de inatención y déficit de atención en el niño

Roselló-Mir (1996) se refiere a la atención bajo el epígrafe de «*la insoportable levedad del concepto...*» (p. 20 y ss.) y es que realmente resulta complicado su abordaje. En nuestro caso realmente lo que más nos interesa es la vertiente aplicada, educativa y clínica, de los problemas de inatención en el niño. Pero para empezar nos encontramos con

un problema crónico: la gran distancia que ha habido siempre entre los trabajos de investigación básica y los de la psicología aplicada. Ruiz-Vargas y Botella (1987, p. 107) critican a los psicopatólogos refiriéndose a la alteraciones atencionales de la esquizofrenia (pero es generalizable a otros trastornos y a otros profesionales más del ámbito infantil) de este modo: *«durante varias décadas han hablado de atención sin disponer de un criterio único para referirse a este constructo y, además, con frecuencia, parecen haber ignorado la evolución de los modelos de atención desarrollados por la psicología experimental»*.

Servera (1999) aborda en la introducción de su trabajo esta misma problemática pero remarcando que en el ámbito infantil tal vez la situación ha sido aún si cabe más grave. El autor entiende que educadores y clínicos diagnostican con mucha frecuencia problemas atencionales en los niños de modo global. Entran en la misma categoría conductas de distracción, de no escuchar, de abandonar la tarea, de falta de concentración, de falta de discriminación, de levantarse de la silla, etc. Y, en la misma línea, como señalan López Soler y García Sevilla (1997), los procesos atencionales y las conductas de atención no siempre son sinónimos: es posible simular gran atención visual y auditiva hacia el profesor y tener el pensamiento a años luz de la clase. Tampoco es del todo adecuado definir la atención en función de la ejecución sobre determinadas tareas, puesto que fácilmente intervienen experiencias previas de aprendizaje y otros procesos psicológicos (memoria, percepción, razonamiento, etc.). En cualquier caso, y pese a esa falta de ligazón entre los modelos experimentales y la práctica clínica, padres y educadores no dejan de referirse a los problemas atencionales de los niños como los más frecuentes, sino los más graves en algunos casos. Manga, Fournier y Navarredonda (1995) señalan que en los estudios epidemiológicos con población normal los maestros describen como inatentos casi a la mitad de los niños y algo más de una cuarta parte de las niñas. Si pasamos al ámbito clínico los problemas atencionales, al menos los relacionados con el trastorno hiperactivo, presentan una prevalencia entre el 3-5% (APA, 1995), e incluso algunos autores señalan que la mitad de los niños que son remitidos para evaluación y/o tratamiento clínico presen-

tan algún tipo de disfunción atencional (Morris y Collier, 1987). Por tanto, parece que ha llegado el momento de abordar esta situación aunando esfuerzos y desarrollando herramientas de evaluación y programas de intervención que deriven de una concepción sólida de lo que es e implica el mecanismo atencional.

De todos modos, antes de abordar la conceptualización y la evaluación nos parece necesario empezar por intentar responder a esta pregunta: ¿qué es realmente un problema atencional en el niño?

López-Soler y García-Sevilla (1997, p. 57) indican que este problema implica *«fallos en los mecanismos de funcionamiento de la atención que producen una falta de adaptación a las exigencias del ambiente, o a nuestras propias exigencias»*. La definición es algo tautológica si no aclaramos a qué mecanismos de funcionamiento nos referimos. Básicamente, aunque posteriormente los ampliaremos, nos referimos a la amplitud atencional, es decir, la cantidad de información que el niño puede procesar al mismo tiempo y a la intensidad atencional, es decir, a la capacidad para mantenerse sobre la tarea, respondiendo a sus demandas. En términos que utilizaremos a menudo a lo largo del capítulo la amplitud define una función de selección o focalización (atendemos a un estímulo objetivo dejando de lado todo lo que en aquel momento es irrelevante) y otra de cambio u oscilación (poder atender a varias tareas o estímulos a la vez), mientras la intensidad se refiere a la capacidad de mantenerse vigilante, alerta o activado.

Un niño puede tener problemas atencionales tanto en la amplitud como en la intensidad, u obviamente en los dos. Pero no es fácil ni determinar en cuál de ellos, ni tan solo si realmente los tiene. La determinación de la presencia o no de un problema atencional en el niño depende de su desarrollo evolutivo, de factores intrínsecos u orgánicos y de factores extrínsecos o ambientales. La dependencia de todo lo que es el mecanismo atencional de esta triple alianza es tan grande que dificulta mucho la tarea del clínico que pretende obrar de un modo objetivo.

Todos los procesos, mecanismos y operaciones cognitivas humanas tienen gran dependencia del desarrollo evolutivo, pero sin duda la atención en grado extremo. Un niño de siete años es capaz de recono-

cer y recordar los nombres (por otra parte nada familiares) de 150 pokémon (una conocida serie de animación infantil), lo cual es indicativo del desarrollo de una capacidad memorística relativamente importante. En cambio no suele resultar muy difícil engañarle en juegos como el trile, descubrir en qué montón de cartas se halla la escogida, o en otros donde «*la vista es la que trabaja*» y debe hacerlo rápido. En este sentido las diferencias entre la capacidad atencional de adultos y niños, o entre niños de diferentes edades, es un hecho ampliamente comprobado. No obstante, y aún disponiendo de una cantidad ingente de información, como explican López-Soler y García-Sevilla (1997, p. 44) no disponemos de un cuerpo de investigaciones que pueda sistematizar detenidamente cuál es el proceso gradual del desarrollo atencional. Hay, eso sí, diversos modelos teóricos que intentan explicarlo (Burack y Enn, 1997; Bermejo, 1987; Vega, 1984); unos creen que existe un desarrollo propio de los mecanismos atencionales independiente del desarrollo de los restantes procesos cognitivos (Lane y Parson, 1982). Para éstos la distraibilidad y el menor control atencional caracterizan la atención infantil frente a la del adulto. Otros creen que el desarrollo atencional depende del de otros procesos, especialmente la percepción (Odom, 1982). Y, finalmente, un tercer grupo sostiene que el desarrollo atencional va unido al desarrollo cognitivo general de la memoria, la inteligencia, el razonamiento, etc. (Hagen y Wilson, 1982). En cualquier caso no estamos en muy buena disposición para responder a preguntas tales como son cuánta información visual o auditiva puede atender un niño a los 5, a los 7 o a los 10 años, cuántas tareas puede realizar simultáneamente o cuánto tiempo puede mantener la atención. En definitiva, si pretendemos detectar un problema atencional en el niño hay que partir de la base que cuánto más pequeño es menor capacidad tiene para atender a la información relevante, más susceptible es a la distracción, tiene menor flexibilidad para orientarla, funciona con menos automatismos y tiene un menor autocontrol del mecanismo, así como una menor capacidad para mantenerla en el tiempo. El problema es que no es fácil cuantificar estas diferencias a lo largo del ciclo evolutivo.

Por si esto no fuera poco, aún si dispusiéramos de un marco evolutivo sólido todavía nos encontraríamos con que la capacidad atencional es altamente fluctuante a causa de determinados factores orgánicos y ambientales. Entre los primeros están muchas enfermedades o problemas de alimentación, así como los déficits sensoriales visuales y auditivos. También podemos incluir las características personales, como el temperamento, determinados rasgos de personalidad o, muy especialmente, la capacidad de autocontrol: la impulsividad motora, los problemas para demorar gratificaciones o la poca tolerancia a la frustración. También algunos estados transitorios como la fatiga, el estrés (la presión) y el sueño alteran fácilmente el funcionamiento atencional. Entre los factores ambientales están las características físicas de los estímulos, la presencia de distractores o estímulos perturbadores, las características de la tarea y, muy especialmente, los aspectos motivacionales y las expectativas. Si bien este último factor también puede depender de los intereses del propio sujeto.

La influencia de los factores evolutivos, orgánicos y externos nos puede llevar a una confusión a la hora de definir un problema atencional en el niño: podemos observar «*conductas de falta de atención*», atribuibles a estos aspectos, y conceptualizarlas erróneamente como «*déficits de atención*». Hay muchos ejemplos en los que se da esta situación y que nosotros en nuestro trabajo hemos podido observar en las aulas:

- Un niño de estilo cognitivo impulsivo aparentemente presenta también un déficit atencional, puesto que no *discrimina bien las figuras*, pero ello es una inferencia tal vez muy arriesgada. La atención focalizada y mantenida que exige la prueba es sólo un mecanismo de base, que puede o no estar alterado, pero el resultado sobre los tests MFF no lo indica: el mecanismo puede funcionar perfectamente pero las estrategias superiores que exige la prueba no, de modo que puede ser impulsivo con independencia de su capacidad atencional.
- El anterior ejemplo puede generalizarse a cualquier tarea de resolución de problemas relativamente compleja (es decir, que la

respuesta no sea inmediata y que exija la puesta en marcha de funciones ejecutivas). No podemos inferir cada vez que un niño presente un pobre rendimiento sobre estas tareas que tiene *un problema atencional*, puesto que esto reduce al absurdo la detección de este tipo de problemas.

- Un niño de 6-7 años que no sabe sumar, o tiene aún muchos problemas para ello, puede que no aguante las diez operaciones de que consta la ficha que ha elaborado el maestro. Las bajas expectativas de autoeficacia y resultado pueden restarle ganas de dedicar más esfuerzo a la tarea y pronto puede buscar otros alicientes hablando a compañeros, canturreando o iniciando una batalla entre el lápiz y la goma.
- Planificar tareas demasiado monótonas, demasiado largas, demasiado complejas en función no sólo de la edad del niño, sino atendiendo a que hay muchas diferencias individuales, y más en la infancia, en los mecanismos atencionales puede provocar multitud de conductas de falta de atención que en absoluto indican déficits en los mecanismos. El cansancio, el aburrimiento, los ritmos circadianos también son factores a tener en cuenta.
- Un niño con trastornos del comportamiento, ya sea por negativismo desafiante o incluso conducta antisocial, puede parecerse inatento porque nunca escucha, hace lo que le viene en gana, rechaza las actividades que suponen esfuerzo cognitivo, etc. pero en absoluto tiene porqué padecer un déficit atencional. Como bien saben los padres y maestros es capaz de atender y aprender con suma facilidad todo aquello que le interesa, de igual modo que puede presentar un bajo rendimiento académico no por falta de ninguna capacidad, sino por esos mismos problemas de comportamiento.

En resumen, lo que queremos expresar es que la mera observación de conductas de falta de atención (distracción, movimiento de pies y manos, cansancio, errores, etc.) no es indicativa de la presencia de un

déficit en el funcionamiento de los mecanismos de amplitud o intensidad atencional. Por tanto, y en parte siguiendo a López-Soler y García-Sevilla (1997), cuando iniciamos un proceso de detección de un déficit atencional infantil en primer lugar hay que analizar el tipo de conductas de falta de atención que padres o maestros han detectado a través de escalas de evaluación, listas de chequeo o registros de observación. En segundo lugar hay que analizar si están presentes variables organizmicas (como, por ejemplo, problemas de vista u oído) o externas (falta de motivación, ambiente poco apropiado, etc.) que puedan ser responsables de dichas conductas. En tercer lugar hay que tener en cuenta la edad del niño y evaluar de algún modo su nivel de desarrollo, con especial atención al desarrollo de la inteligencia o a las capacidades más específicas de procesamiento de la información (también pueden incluirse pruebas de percepción, de lateralidad, de orientación temporo-espacial, etc.). Por último, es necesario aplicar pruebas atencionales específicas de amplitud (selección/focalización y cambio atencional) y/o intensidad (capacidad de mantenimiento) con índices psicométricos suficientes como para poder valorar la puntuación del niño en relación a lo que sería de esperar. Con toda esta información en la mano sin duda tenemos muchas más posibilidades de acertar en nuestro diagnóstico de déficit de atención que, por otra parte, puede integrarse en un síndrome más complejo como el trastorno por déficit de atención con hiperactividad (en adultos con más facilidad se le detecta en múltiples psicopatologías) o como una dificultad de aprendizaje.

Para acabar, y en relación a las pruebas objetivas de evaluación de la atención, cabe resaltar, en primer lugar, que no son muchas y, en segundo lugar, que muchas veces se hallan contaminadas por otros procesos o habilidades cognitivas, especialmente la percepción, pero también con factores de aprendizaje y razonamiento. Por ejemplo, los conocidos tests de Caras (percepción de diferencias), Cuadrados de letras, SIT (Situación 1), FI (Formas idénticas), Tests de Figuras Enmascaradas o la prueba del color de Stroop se utilizan en mayor o menor grado como medidas atencionales, si bien en todos ellos parece evidente la presencia de la contaminación a la que antes aludíamos.

Dejando de lado las múltiples pruebas experimentales desarrolladas, que en realidad abarcan todas las funciones atencionales detectadas, en la práctica las tres medidas atencionales más utilizadas, pero no exentas de polémica, son los tareas por ordenador de vigilancia o de ejecución continua, las tareas de «tachado», donde el niño resigue un conjunto de estímulos similares impresos en una hoja de papel y va tachando uno determinado (es el caso del clásico Toulouse-Pieron o de la más reciente EMAV de García-Pérez y Magaz-Lago, 1999), que realmente son estrictamente hablando de ejecución continua, y el factor *libre de distractibilidad* del WISC-R compuesto por las pruebas de aritmética, dígitos y claves. En todos estos casos al menos los niños hiperactivos normalmente rinden peor que los niños normales, y también es muy habitual que niños con déficit de atención, sin estar asociado a problemas de impulsividad-sobreactividad, también muestren mayores problemas (véase Servera, Bornas y Moreno, 2001).

3.2. El funcionamiento atencional

3.2.1 Definición y concepto

La atención puede entenderse de muy diversas maneras. Puede ser un proceso cognitivo, un estado de la mente, una volición o incluso un mecanismo. Pero Ruiz-Vargas y Botella (1987, p. 95) hacen especial énfasis, en contra de una posición muy extendida, que la atención «no es un proceso cognitivo» (p. 95), sino la actividad mediadora que participa en todos los procesos cognitivos. García Sevilla (1997, p. 14), en una línea similar, define la atención como «el mecanismo implicado directamente en la activación y el funcionamiento de los procesos y/u operaciones de selección, distribución y mantenimiento de la actividad psicológica». La capacidad de funcionamiento de este mecanismo es relativamente innato, pero se complementa con un segundo aspecto fundamental: la habilidad o estrategia aprendida. Las personas diferimos en nuestra capacidad para concentrarnos en una cosa y abstraernos de

otras para atender a dos o más eventos a la vez y para mantener durante una cierto tiempo la activación necesaria para el correcto rendimiento, y esas diferencias se concretan en distintas estrategias atencionales, que en realidad son un tipo de habilidad cognitiva aprendida que se van desarrollando y combinando con otras de orden superior (la metatención).

Por supuesto resulta casi imposible evaluar la atención de manera separada de otros procesos o mecanismos como, por ejemplo, la percepción, que durante mucho tiempo parece que ejerció de «hermana mayor». También es muy difícil separar la atención de la memoria de trabajo que, en general, se han integrado conjuntamente en los modelos explicativos de la organización y recuperación de la información. En ambos casos la atención actuaba sólo de puerta de entrada de los procesos cognitivos: su función era de selección inicial. Actualmente, sin embargo, la mayoría de autores le conceden una capacidad de funcionamiento en vertical, de articulación y control de todos los procesos cognitivos (García-Sevilla, 1997; Rosselló-Mir, 1996; Ruiz-Vargas y Bottella, 1997). Las características más importantes de este funcionamiento de la atención son cuatro: amplitud, intensidad, oscilamiento y control. Tal vez, en los últimos tiempos, el control, que hace referencia a la distinción entre un funcionamiento automatizado o libre de la atención y otro controlado o dirigido, ha acabado constituyéndose en la característica fundamental.

La atención se manifiesta, y por tanto es susceptible de evaluación, a través de respuestas generadas por el sistema nervioso (fisiológicas y motoras), por el rendimiento cognitivo o por la experiencia subjetiva de «estar atento». García-Sevilla (1997) y Rosselló-Mir describen las primeras y las últimas, pero a nosotros nos interesa destacar fundamentalmente las segundas. Hay multitud de tareas/respuesta que nos sirven para valorar la cantidad y la calidad de la atención de las personas, como, por ejemplo, el tiempo de reacción, la discriminación, la identificación, el recuerdo, el reconocimiento o la búsqueda estimular.

Los determinantes de la actividad atencional son múltiples, aunque, como hemos visto, tradicionalmente se han dividido en extrínsecos (re-

lativos a las características estimulares) e intrínsecos (relativos al estado del sujeto), García-Sevilla (1997) atendiendo al hecho probado de que ambos actúan a menudo de manera interrelacionada opta por una clasificación más abierta centrada en las características de los objetos, el nivel de activación fisiológica, los intereses y las expectativas, y los estados transitorios. Veamos brevemente cómo actúan.

La preocupación de qué dimensiones físicas de los objetos captan mejor nuestra atención es, como se comenta a menudo, una obviedad en una economía de libre mercado. El interés que va mucho más allá de los psicólogos experimentales y, por tanto, se le ha dedicado gran esfuerzo, siendo tal vez los niños los sujetos mejor estudiados. Hoy en día sabemos que captan nuestra atención los estímulos de mayor tamaño, los colocados en las partes superiores, más a la izquierda que a la derecha, los de color frente a los de blanco y negro, los de mayor intensidad, los de mayor movimiento, los de mayor complejidad, los de mayor relevancia (inducida por haber adquirido anteriormente una significación especial para el sujeto), los más novedosos (la novedad no sólo se consigue por el efecto sorpresa sino simplemente alterando cualquiera de las anteriores características nombradas). Claro está que estas características pueden variar su potencial de atracción actuando de manera interdependiente o combinada y que, en el fondo, la intensidad, la novedad y especialmente la relevancia son los elementos fundamentales. Éste último por cierto trasciende lo que es meramente la característica física para relacionarla con los intereses, las motivaciones o los aprendizajes previos del sujeto.

El nivel de activación o *arousal*, esto es, el nivel de receptividad y responsividad que el sistema nervioso posee en un determinado momento, es clave para entender el funcionamiento atencional. La activación interviene especialmente en los procesos emocionales y los mecanismos atencionales. En el caso que nos ocupa, «*estar activado*» es imprescindible para garantizar capacidad de concentración y mantenimiento atencional. Por supuesto no sólo puede ser perjudicial la falta de dicha activación, sino también el exceso. Así, la eficacia en la ejecución de una tarea, en relación al arousal, suele presentar forma de U in-

vertida: a medida que aumenta nuestra activación aumenta nuestro rendimiento hasta llegar al punto óptimo, por encima de este punto la ejecución decrece por efecto de sobreactivación o cansancio. De todos modos el tipo de tarea matiza bastante la forma de esta curva: ante tareas sencillas la ejecución es mejor en niveles altos de activación, sin embargo ante tareas complejas de entrada suelen ser mejores niveles de activación más bajos. Por otra parte, otra línea de trabajo muy interesante ha sido el estudio de la variabilidad del nivel de activación en función de las horas del día; los ritmos biológicos circadianos tienen influencia directa sobre la atención, aunque de forma variable en función de diferencias individuales. Por ejemplo, se ha supuesto que las tareas atencionales perceptivo-motoras se van realizando mejor conforme avanza el día, mientras las memorísticas funcionan mejor al principio. Por otra parte, dentro de las horas del día se han establecido momentos puntuales de máxima y mínima atención: las primeras horas de la mañana y primeras de la noche ocupan los mejores lugares, mientras a media mañana o tras las comidas cabe esperar peor rendimiento. En cualquier caso los agentes estresores o las contingencias del medio, además de las propias diferencias en estrategias atencionales, también juegan un papel fundamental, lo que sin duda complica aún más el tema.

En cuanto a los intereses y las expectativas, factores motivacionales en definitiva, también cabe considerarlos clave para el análisis de la función atencional. Respecto a los intereses es lógica la confirmación de que aquellos estímulos que tienen algún tipo de interés para la persona se perciben antes y mejor. En el tema de las expectativas se han desarrollado paradigmas experimentales muy diversos para analizar el efecto del llamado *set atencional*, esto es, cómo se ve alterada la ejecución de la persona cuando previamente se le da algún tipo de información sobre los estímulos que recibirá. Esta información es conocida como el *efecto priming* o efecto de facilitación/preparación (Rosselló-Mir, 1996). Su influencia sobre el rendimiento del sujeto puede ser positiva (si realmente es facilitador) o negativa (si, por contra, es incongruente o inesperado para el sujeto).

Finalmente los estados transitorios a los que aludíamos como determinantes de la atención se refieren a estados del propio sujeto. Entre los más destacados están la fatiga, el estrés, el efecto de ciertas drogas y psicofármacos y el sueño. Lógicamente todos los determinantes citados, más otros en lo que nos hemos entrado directamente, pueden funcionar por separado o de manera interactiva.

La clasificación de los distintos tipos de atención tampoco escapa de esta complejidad, puesto que los criterios que se pueden seguir son múltiples. Rosselló-Mir (1996) destaca los siguientes tipos de atención: voluntaria versus involuntaria, abierta versus encubierta, dividida versus selectiva, visual versus auditiva y «atención» versus «vigilancia» (o atención selectiva versus atención sostenida).

La atención voluntaria se diferencia de la involuntaria en función del grado de control del mecanismo, al estilo de lo que ya comentamos anteriormente al hablar de las características de la atención en general. Así, nuestra atención puede ser captada de manera refleja por determinados estímulos (el llamado control *bottom-up*), o bien la podemos dirigir según nuestra voluntad a los estímulos deseados (control *top-down*). La atención abierta es la que es susceptible de observación: hay presencia de respuestas motoras y fisiológicas (el reflejo de orientación). Por contra la atención encubierta no permite una observación externa de la acción del sujeto. La atención selectiva o focalizada hace referencia a la capacidad de procesamiento diferencial de las fuentes estimulares, mientras la dividida hace referencia a los recursos, la capacidad, el esfuerzo, en definitiva el procesamiento en paralelo de información. Obviamente la diferenciación entre atención visual y auditiva hace referencia al canal sensorial que se estudie, y esta opción ha tenido importantes repercusiones sobre los modelos teóricos y sobre el tipo de atención estudiada: los modelos de *cuello de botella*, como el de filtro de Broadbent y otros centrados en la selectividad atencional, van tan estrechamente ligados a la modalidad auditiva que se hace difícil pensar en ellos en la visual. Por contra, los modelos de procesamiento en paralelo, de cambio y capacidad atencional, se han desarrollado mucho más en la modalidad visual.

La última distinción propuesta por Rosselló-Mir (1996, pp. 27-28) es más controvertida. El autor propone una necesaria distinción entre lo que es «*atención*» y «*vigilancia*», términos confusos en el lenguaje común y hasta en la práctica científica. En concreto propone que nos refiramos a «*atención*» siempre que hablemos de los mecanismos de selectividad/focalización y de división del mecanismo atencional y únicamente utilizemos «*vigilancia*» para referirnos al mecanismo de mantenimiento. En el primer caso lo que se estudian son los déficits ejecutivos derivados de la sobrecarga de estimulación o de actividad en un entorno en donde el objetivo es localizar un estímulo enmascarado o cambiante (focalización y/o cambio atencional). Mientras en el segundo caso lo que se estudia son los déficits ejecutivos derivados de la distracción (sin necesidad de estimulación externa) o el cansancio.

Estamos de acuerdo básicamente con la propuesta de Rosselló-Mir y sin duda sería conveniente que esta terminología se asentase en el ámbito clínico, ayudando a desterrar la confusión que se produce actualmente cuando un niño recibe el diagnóstico de «*déficit atencional*» a través de múltiples sistemas de evaluación y significados, a veces incluso desconocidos por el propio clínico.

3.2.2 Los tipos de atención: selectiva, dividida y sostenida

Como decíamos, en la práctica, los tipos de atención más trabajados son la selectiva, la dividida (estas dos muchas veces de manera conjunta) y la sostenida. Brevemente la atención selectiva, también llamada focalizada, hace referencia a la capacidad para atender a cierta estimulación relevante dejando de lado la irrelevante. Está muy relacionada con la concentración (aunque este concepto abarcaría elementos de los otros tipos de atención, así como otros procesos cognitivos), con la presencia de estímulos distractores y con la capacidad de anticipación de los elementos que serán relevantes. Por su parte, la atención dividida hace referencia a una doble capacidad: la de cambiar rápidamente de tarea en función de las demandas y la de atender simultáneamente a dos o más tareas. En el primer caso se habla de flexibilización u oscila-

miento atencional, mientras en el segundo se habla de distribución de recursos y de efectos de interferencia. Por último, la atención sostenida hace referencia a la capacidad para mantener un cierto estado de alerta o de activación durante períodos relativamente largos y ante tareas repetitivas o monótonas. En este caso, como veremos, se puede hacer una distinción entre tareas que exigen rendimiento continuo (durante todo el proceso el sujeto debe realizar algún tipo de actividad simple, por ejemplo, aceptar o rechazar la presencia de un estímulo objetivo) o tareas de vigilancia (el sujeto debe estar atento a un determinado estímulo que aparece muy de vez en cuando, y responder únicamente ante su presencia).

La definición de estos distintos tipos de atención guarda mucha relación con la evolución de los principales modelos explicativos del mecanismo atencional. No es el momento de entrar en detalle en ellos pero sí de establecer algunas características que nos ayuden a una mejor comprensión de los tipos mencionados y las distintas formas de evaluar la atención (véase para una ampliación, por ejemplo, Baños y Belloch, 1995, Munar, Rosselló-Mir y Sánchez-Cabaco, 1999, Rosselló-Mir, 1996 ó García-Sevilla, 1997).

De forma tradicional los modelos explicativos de la atención se engloban en dos categorías: estructurales (o de filtro) y de capacidad (o de recursos). Los modelos estructurales enfatizan sobre todo la *selección* (Baños y Belloch, 1995). Se postula la existencia de mecanismos o estructuras que impiden un procesamiento simultáneo de la información, siendo la atención la encargada de regular la entrada serial de los *inputs*. La teoría del filtro de Broadbent (1958) inició este tipo de modelos basándose en los experimentos con tareas de escucha dicótica y de amplitud dividida. Según esta teoría, los estímulos deben pasar por un canal de capacidad de limitada (*un cuello de botella*) de manera que necesariamente hay que hacer una selección. Este filtraje se basa en las características físicas del estímulo, es decir, se haría de abajo/arriba (*bottom-up*), si bien el sujeto podría anticipadamente programar ciertas prioridades.

Las tareas de escucha dicótica se utilizan como una medida de atención selectiva o dividida. En ellas, el sujeto escucha a través de unos au-

riculares voces distintas (por ejemplo, una masculina y otra femenina, una por cada oído) que le transmiten algún tipo de información (dígitos, letras, palabras o prosa). Al sujeto se le pide o bien que atienda a un solo mensaje o bien que atienda a elementos de los dos. Existen dos variedades en este tipo de tareas: la técnica del sombreado exige del sujeto que vaya repitiendo, resiguiendo, un determinado mensaje mientras por el otro oído se le suministra un mensaje irrelevante al que no debe atender. Es, por tanto, una técnica de evaluación de la selectividad atencional. Por su parte, la técnica de amplitud dividida exige del sujeto que atienda a dos mensajes que normalmente no son simultáneos, sino que se presentan alternativamente y de manera muy rápida. La información suelen ser dígitos y en este caso deben repetirse al haberse escuchado, con lo cual hay un componente muy importante de memoria de trabajo. Esta técnica sirve fundamentalmente para estudiar el cambio y la amplitud atencional.

Con el paso del tiempo las investigaciones demostraron que las teorías del filtro presentaban bastantes limitaciones (véase García-Sevilla, 1997 o Rosselló-Mir, 1996). Las distintas reformulaciones han llegado a aceptar, por ejemplo, que a parte del sistema de filtraje, básicamente físico, existe otro sistema de selección de respuesta básicamente semántico que relativiza la función de todo/nada del canal de capacidad.

Los modelos de capacidad o de recursos tienden a obviar el problema de la selectividad en principio porque ésta es obligada ante un sistema que, por definición, debe ser limitado. Conciben el funcionamiento de la actividad cognitiva como un procesador central inespecífico y limitado, cuyo buen funcionamiento consiste en una buena administración de los recursos en función de las demandas internas del sujeto y externas de las tareas. El paradigma de investigación pasa, pues, de las tareas de selectividad a las tareas de vigilancia. El modelo de Kahneman (1973) abrirá las puertas a otros muchos centrados ahora en valorar la ejecución atencional en función de la similitud de las tareas, su grado de dificultad y la práctica previa del sujeto. Las tareas de vigilancia y de ejecución continua son objeto de la parte experimental del presente trabajo y es por ello que las revisaremos con más profusión más

adelante. Ahora es importante destacar como a partir de estos modelos de recursos limitados otros conceptos van apareciendo en el ámbito del estudio de la atención, como la noción de preparatoriedad, es decir, la influencia que tiene la información previa sobre el rendimiento atencional del sujeto, y la distinción entre procesamiento controlado y automático de Shiffrin y Schneider (1977).

Entre los paradigmas de investigación que envuelven estos nuevos conceptos y los principios de la distribución de recursos cabe destacar el de la doble tarea, el de Stroop, el de la búsqueda visual y el del *set* atencional. En el paradigma de la doble tarea básicamente lo que hay que hacer es realizar simultáneamente dos tareas. Normalmente se produce un efecto de interferencia que provoca el deterioro del rendimiento en una de ellas, o en las dos, medible a partir de una línea base en donde cada tarea se ha resuelto por separado. En el efecto Stroop, o prueba de interferencia color-palabra, se le presenta al sujeto palabras impresas en distintos colores, siendo su labor el nombrar el color de la tinta en que se han sido escritas. En la condición facilitadora las palabras son nombres de colores y precisamente coinciden con el color de la tinta en el que se han escrito; en la condición de interferencia las palabras siguen siendo nombres de colores pero ahora no coinciden con el color de la tinta; finalmente en la condición de control las palabras ya no son nombres colores (son objetos familiares). Los efectos de interferencia se enmarcan dentro de lo que es el cambio y la distribución atencional, si bien el efecto Stroop también hace referencia a la flexibilidad para pasar del automatismo al control atencional.

En el paradigma de la búsqueda visual la tarea consiste en buscar, comparar y reconocer determinados estímulos, memorizados previamente, y que ahora se presentan en conjuntos estimulares que pueden o no contenerlos. Las estrategias de exploración y búsqueda que se elicitan guardan relación con la atención selectiva y con los procesos de automatización. El paradigma del *set* atencional consiste en que de algún modo se informa al sujeto de la inminente aparición del estímulo objetivo o *target*, al que debe responder. En la variante del *priming* o

preparación un estímulo antecede a otro, de forma que el primero afecta a la ejecución del segundo. En la variante de *costes y beneficios* los estímulos antecedentes pueden proporcionar información válida, no válida o neutra sobre la aparición del estímulo objetivo. La anticipación o preparatoriedad guarda relación tanto con la selección como con el cambio atencional.

En el trabajo de Servera (1999) se revisan en detalle las implicaciones en psicopatología de la atención selectiva, dividida y sostenida en adultos y niños. En lo que respecta al ámbito infantil los problemas de atención se han centrado en las dificultades de aprendizaje, y muy especialmente en los trastornos por déficit de atención con hiperactividad. Como veremos en el siguiente apartado, la atención sostenida, a través de tareas de vigilancia y ejecución continua, han ocupado el lugar central. Sin embargo, uno de los aspectos diferenciales más importante de la investigación con niños respecto a la de los adultos es, como ya avanzamos previamente, el tema del desarrollo evolutivo.

Tal vez los datos más fiables son los de los recién nacidos, en estos casos está muy estudiado el reflejo de orientación, las preferencias por ciertos sonidos, la atención visual captada por estímulos novedosos e intensos, etc. Llama la atención un poderoso efecto de habituación que hace, incluso en el niño cercano a los 2 años, que los estímulos pierdan muy rápidamente el interés. Precisamente se considera que es en esa edad cuando empiezan a diferenciarse los tipos de atención selectiva, dividida y sostenida. En la selectividad cabe destacar que ahora las características de los estímulos van perdiendo poder de captación porque el niño va ejerciendo un mayor control sobre el mecanismo. Hasta los 6 o 7 años este control será relativamente débil. En cuanto a la oscilación o cambio atencional los niños son mucho más lentos que los adultos, igualmente la presencia de estímulos distractores les afecta mucho más (son especialmente relevantes los datos con el efecto Stroop). Las tareas de amplitud dividida también diferencian claramente a niños menores de 7 años de los mayores, aunque incluso hasta los 17 años puede mejorar el rendimiento.

Por su parte, se considera que el desarrollo de la atención sostenida es más lento y posterior al de la atención selectiva, aunque tal vez no sea un efecto directo del distinto grado de maduración de estos mecanismos, sino de la mayor susceptibilidad a la fatiga, al cansancio o a los estímulos externos de los niños más pequeños. García-Sevilla (1997) concluye que un niño de 2 años puede atender eficazmente hasta siete minutos, llegando a los 14 minutos a los 5 años y es alrededor de los 8-9 años cuando se produce una mejora crítica en las tareas de vigilancia. En nuestra experiencia, y como veremos en la segunda parte del presente trabajo, estas cifras son relativas: poder atender más o menos minutos depende mucho del tipo de tarea. Es posible que los datos sean acertados si hablamos de atender a cuentos, programas de televisión, o incluso a un maestro hábil enseñando (y por ende, hábil captando la atención del niño), pero son más dudosos si hablamos realmente de tareas de vigilancia, monótonas y repetitivas, sin posibilidad por parte del niño de recibir feedback inmediato. En estos casos nuestra experiencia demuestra que tareas de vigilancia entre 7 y 8 minutos son suficientes para elicitar diferencias individuales importantes en niños entre 6-7 y 9-10 años, aunque tal vez estaría por ver si esas diferencias globales también se acompañan del habitual menoscabo atencional que debe generarse al medir la atención sostenida. En cualquier caso veamos a continuación las características de la medida de este mecanismo, que resulta central en nuestra investigación.

3.3. La evaluación de la atención sostenida: las tareas de ejecución continua y las tareas de vigilancia

Como hemos venido explicando, en el ámbito infantil la atención sostenida ha tenido cierta prioridad en los trabajos aplicados, tanto en lo que respecta a psicopatología como al campo de las dificultades de aprendizaje. Como quiera que una de las alteraciones comportamentales infantiles más importantes, si exceptuamos las deficiencias, como

son los trastornos por déficit de atención con hiperactividad (TDAH) se han relacionado a menudo con déficits en la atención sostenida, no es difícil entender esta situación.

Curiosamente, aunque nos detendremos más en el tema al caracterizar los problemas del niño con déficit de atención con y sin hiperactividad, el sentido común nos hace pensar que la distraibilidad y los problemas de selección estimular deben estar en la base de los problemas de los niños con TDAH. Sin embargo, múltiples estudios (con pocas excepciones) han podido observar que la presencia de distractores no supone un problema especialmente grave para que los niños con TDAH ejerzan un adecuado control (véase las revisiones de Van de Meere, 1996, p. 123; Bornas, Servera y Moreno, 2001; Servera, 1999). En cambio sí hay un cierto acuerdo en que estos niños tienen problemas importantes de atención sostenida.

Las dos pruebas de evaluación más utilizadas han sido las tareas de ejecución continua (de ahora en adelante TEC, traducción de sus siglas en inglés, CPT —Continuous Performance Tests—) y las tareas de vigilancia (de ahora en adelante TVI). En los trabajos aplicados los autores no han tenido mucho cuidado en diferenciar ambas pruebas, y a veces las TEC incluyen una y otra posibilidad. Pero como hemos visto derivan de paradigmas experimentales diferentes: **las TEC son tareas en donde el niño debe dar algún tipo de respuesta (simple) continuada durante tiempos prologados de tiempo, mientras las TVI son tareas en las que el niño debe permanecer alerta y sólo debe responder de vez en cuando ante la presencia de un determinado estímulo.** El hecho de que a menudo se use el término «*tarea de ejecución continua*» como sinónimo de ambos tampoco es de extrañar, puesto que con el mismo formato de actividad, cambiando sólo los parámetros, ya es posible pasar de uno a otro modelo experimental. En realidad incluso cabe la posibilidad de cambiar a distintos paradigmas de investigación atencional. Veamos un ejemplo.

En su estructura más actual las pruebas de evaluación de atención sostenida consisten en colocar al niño delante de la pantalla del ordenador con la instrucción de que permanezca atento a la aparición de

una serie de estímulos (letras, números, dibujos, etc.) y cada vez que salga el objetivo (habitualmente conocido como *target*) debe apretar, por ejemplo, la barra de espacio del aparato. Si lo que proyectamos son números del 0 al 9 y definimos el 6 como *target* cabe hablar de una TVI si su tasa de aparición es, por poner un límite, inferior al 10%, siempre y cuando la tarea sea lo suficientemente larga. La misma prueba se podría convertir en una TEC si obligamos al niño a responder ante cada estímulo: con una tecla indicará que ha visto el *target* y con otra que no. En este caso debe estar respondiendo durante todo el tiempo que dure la tarea, pero seguimos en el paradigma de atención sostenida fundamentalmente porque la tasa de aparición del *target* es muy baja. Si con las mismas condiciones aumentamos esta tasa hasta el 50% ya sería posible hablar de la influencia de un mecanismo de atención selectiva: en realidad ahora se trata de focalizar la atención sobre un estímulo dejando de lado todos los demás, que funcionan como distractores. Pero aún podríamos realizar otras modificaciones para acercarnos, por ejemplo, a la atención dividida. Para ello bastaría introducir una pequeña ventanilla en la pantalla que indicara en cada momento cuál es el *target* vigente (que sería variable), así el niño debería atender a dos tareas relacionadas casi a la vez. Si nos interesa evaluar la capacidad de control del mecanismo atencional podemos introducir de vez en cuando una señal de stop (una luz roja ante la cual no se debe de responder nunca) y si queremos provocar interferencia le podemos colocar unos auriculares y que vaya oyendo números al mismo tiempo que se proyectan, sin que normalmente se dé coincidencia entre el que ve y el que oye.

En cualquier caso nuestro interés está especialmente en la atención sostenida y a ella hemos dedicado nuestros esfuerzos. Cabe matizar simplemente que en trabajos anteriores y en el mismo proyecto de investigación que presentamos al Centro de Investigación y Documentación Educativa, utilizábamos en su acepción amplia el término «*tarea de ejecución continua*». En realidad otros colegas y nuestra propia experiencia nos aconseja hablar más en propiedad y utilizar el término «*tarea de vigilancia*», que en este trabajo aparecerá definitivamente.

3.3.1 Características generales y tipos de medida en las tareas de vigilancia y ejecución continua

Una vez vistas las diferencias entre lo que son las TEC y las TVI, pasaremos a revisar sus características generales, que en realidad, y aunque aquí ya nos centremos fundamentalmente en las segundas, son prácticamente compartidas.

La atención sostenida se define como *«la actividad que pone en marcha los procesos y/o mecanismos por los cuales el organismo es capaz de mantener el foco atencional y permanecer alerta ante la presencia de determinados estímulos durante períodos de tiempo relativamente largos»* (García-Sevilla, 1997, p. 140). Estas tareas requieren esfuerzo y acaban provocando siempre un deterioro en la ejecución. Este deterioro puede medirse de dos maneras: en primer lugar por la función de decremento o menoscabo atencional, esto es, por el declive progresivo en la ejecución, que en adultos suele situarse entre los 20 y 35 minutos. Normalmente esta es la medida más utilizada por los investigadores básicos, los que se centran en el desarrollo de teorías puras del mecanismo atencional. En segundo lugar está la función del nivel o capacidad de vigilancia, es decir, el rendimiento final y no a lo largo del tiempo. Esta segunda medida se aproxima más a nuestro trabajo, y a lo que ha sido la psicología aplicada en general: el objetivo es determinar diferencias entre sujetos en su nivel global de vigilancia, y observar si ello repercute en otras variables, que en el caso de los niños suelen ser de aprendizaje o comportamiento.

Históricamente se considera que los estudios de atención sostenida se iniciaron con los trabajos de Norman Mackworth, durante y poco después de la Segunda Guerra Mundial. El paso de estos estudios al ámbito de la psicopatología, sufrió un fuerte impulso con las tareas de ejecución continua para adultos en los años cincuenta (Rosvold y cols., 1956). El impulso en el ámbito infantil se debe a los trabajos de Virginia Douglas (1982, 1983) a finales de los setenta y principios de los ochenta que sirvieron para definir el término *«trastorno por déficit atencional»* en la infancia, y para destacar que era la atención sostenida, y no otros ti-

pos de atención, los más implicados en los problemas de aprendizaje y de desarrollo de los niños. Aunque la cuestión es algo polémica, la mayor lentitud en el desarrollo del mantenimiento de la atención frente a otras funciones (ya comentada anteriormente) y la idea de que su nivel de intensidad influye en cualquier actividad cognitiva del niño (con muchos menos recursos compensatorios que el adulto) hacen de su estudio una prioridad.

Las tareas de ejecución continua y de vigilancia que sirven para medir la atención sostenida actualmente suelen ser por ordenador. En su versión actual, como ya hemos explicado, se suele colocar al niño ante la pantalla del ordenador y se le pide que responda ante la presencia de un determinado estímulo, dejando de lado otros, durante períodos de tiempo largos. La prueba más conocida en el ámbito infantil es el *Conners' Continuous Performance Test* (CPT-II, que puede consultarse en internet en la siguiente dirección: <http://www.mhs.com/onlineCat/product.asp?productID=CPT-WIN>) con una larga tradición en Estados Unidos, pero menos utilizada en otros países, tanto por los problemas de adaptación y baremación de las muestras como por su relativamente elevado alto coste. En cualquier casos, las principales características técnicas que se controlan en cualquier tarea de vigilancia o de ejecución continua son las siguientes:

- *La duración de la prueba:* se trata de pruebas de larga duración, que en adultos suelen estar sobre los 30 minutos, y a veces si se trabaja con TVI pueden durar horas. En niños no hay demasiados datos sobre cuál debe ser esta duración, aunque muchas veces se está alrededor de 15 minutos, aunque los trabajos suelen ser siempre con niños de 10 o más años.
- *Tipo de estímulo a utilizar:* normalmente son letras, números o formas simples. El objetivo es atenuar al máximo la influencia de factores de aprendizaje, de memoria, de razonamiento, etc. Es posible, aunque menos frecuente, trabajar la atención sostenida con estímulos auditivos (normalmente se rinde mejor que con los visuales). El número total de estímulos que se presentan es muy variable.

- *El estímulo objetivo, la señal o target:* se define un estímulo X, o un par de estímulos AX, ante los cuáles el niño debe efectuar algún tipo de respuesta motora, normalmente apretar la barra espaciadora del teclado.
- *El intervalo de presentación del estímulo:* el tiempo que está el estímulo visible también es variable, pero suele estar entre 200 milisegundos hasta el segundo completo.
- *El intervalo interestímular:* el tiempo que transcurre entre la presentación de cada estímulo y que puede ser extraordinariamente variable. En el paradigma de vigilancia pura pueden pasar minutos entre estímulos, mientras en el de ejecución continua pueden ser milisegundos.
- *La razón del evento o ritmo de estimulación:* la tasa de presentación del target es variable, pero normalmente baja. García-Sevilla (1997, p. 149) divide entre tareas en las que el target se presenta menos de 24 veces por minuto y tareas en las que está presente más de 24 veces. Las primeras serían las más propias del paradigma de vigilancia.
- *Otras características:* la carga de la señal, los múltiples canales, el tipo de discriminación, la previsibilidad del target, el conocimiento de resultados, etc. son otras variables que el experimentador puede manipular.

En este tipo de pruebas, como mínimo, se pueden obtener las siguientes medidas:

- *Aciertos y/o Errores de omisión:* el número total de aciertos o, en su defecto, el número de targets que el niño ha dejado de señalar.
- *Tiempo de reacción a los aciertos:* se mide, en milisegundos, la latencia media de respuesta del sujeto en la identificación de los targets.

- *Errores de comisión o falsas alarmas*: el número total de veces que el sujeto ha creído identificar un target cuando realmente no lo era.
- *Tiempo de reacción a las falsas alarmas*: se mide, en milisegundos, la latencia media de respuesta del sujeto en las falsas alarmas.

En la práctica, cuando el interés está más en la función del nivel o capacidad de vigilancia que en la función de menoscabo atencional, las dos puntuaciones más interesantes son la de aciertos y la de falsas alarmas, puesto que, salvo excepciones, es difícil encontrar diferencias en variables educativas o comportamentales sujetas al tiempo de reacción. En realidad, cuando la prueba se ha aplicado en condiciones óptimas, se supone que si un niño, en función de su edad y su nivel de desarrollo, comete más errores de omisión de los esperados presenta una «*disfunción atencional*» que cabe atribuir a un exceso de distraibilidad (más bien como efecto de dispersión de la actividad cognitiva y no por presencia de distractores) o a un exceso de lapsus atencionales. Estos lapsus, por otra parte necesarios en nuestro funcionamiento cognitivo, pueden ser cortos y transitorios (cambios fásicos de la intensidad atencional) o largos y relativamente permanentes (cambios tónicos). Los primeros tienen que afectar mínimamente a la actividad puesto que se recuperan en seguida, mientras los segundos, al menos en adultos, no deberían aparecer hasta aproximadamente los 30 minutos de trabajo en tareas de vigilancia (el sujeto experimenta un estado subjetivo de total aburrimiento y sensación de que no puede más). Por su parte, la naturaleza de los errores de comisión ha generado mayor polémica: Barkley (1990; 1997), y otros autores, defienden la posibilidad de que refieran fundamentalmente un problema de impulsividad motora, es decir un trastorno del bajo control relacionado con una disfunción en los sistemas de activación/inhibición cortical.

En realidad los trabajos de Barkley, y otros muchos investigadores clínicos infantiles, se han dedicado fundamentalmente a observar el rendimiento de niños con trastornos por déficit de atención con hiperactividad (TDAH) sobre las tareas de vigilancia y ejecución continua. A parte de la idea más extendida de que, en su conjunto, estos niños rin-

den peor que niños control o niños con otro diagnóstico clínico, algunos autores han profundizado mucho más en la relación entre la atención sostenida y el TDAH.

El grupo de Van der Meere (véase, por ejemplo, Van der Meere, 1996) ha publicado una cantidad importante de trabajos en los que ha relacionado las tareas de vigilancia con las hipótesis explicativas del TDAH basadas en la baja activación cortical o arousal. Ambos términos se suelen usar de manera sinónima, y además responden a un concepto unitario de los estados de vigilancia, alerta y responsabilidad del sistema nervioso central, pero Van der Meere se suma a los autores que han distinguido ambos términos: arousal hace referencia al efecto de alerta de los sistemas sensoriales, mientras la activación hace referencia al control de la preparación de la respuesta motora. El arousal se localiza en el área límbico-frontal, bajo la influencia primordial de los neurotransmisores noradrenalina y serotonina. El sistema de activación se localiza en los ganglios basales, bajo la influencia fundamental de la dopamina y la acetilcolina.

Algunos autores, como Sanders (1983), han postulado la existencia de un tercer sistema, llamado sistema de esfuerzo (*effort system*), que a su vez depende de un complejo sistema de evaluación. Este último sistema es el encargado de explorar los estados de arousal y activación del sujeto, si alguno de ellos está por debajo del nivel óptimo el sistema de esfuerzo intenta ejercer la compensación (el hipocampo es el área cortical más implicada). El mecanismo de esfuerzo se halla influido por factores motivacionales, como el conocimiento de los resultados, la presencia o ausencia del experimentador (o un adulto relevante para el niño mientras él realiza la tarea), los reforzadores, etc. De acuerdo con el modelo de Sanders los problemas de los niños con TDAH podrían estar o en el procesamiento inicial de la información (el mecanismo de los sistemas arousal/esfuerzo), o en el control de respuesta más avanzado (el mecanismo de los sistemas activación/esfuerzo), o bien en los dos. El grupo de Van der Meere, en contra de las ideas más extendidas, hipotetizó ya en los años ochenta que el déficit estaría más en el canal motor, es decir, en los sistemas de activación/esfuerzo. Para empezar ello podría ser una buena explicación a una cuestión que comentamos

en el capítulo dedicado a la impulsividad: los niños hiperactivos tendían a mostrarse más bien lentos-inexactos y no impulsivos sobre los tests MFF. Para Van der Meere ello podría ser una prueba favorable a su idea.

El grupo de este autor puso a prueba su hipótesis utilizando TEC y TVI. En primer lugar Van der Meere no duda que el rendimiento global de los niños hiperactivos sobre estas tareas sea siempre inferior al de niños control. Pero lo importante para determinar la presencia de un déficit en atención sostenida es, para el autor, el observar la función de decremento de la ejecución a lo largo de la prueba. En función del modelo de Sanders, Van der Meere hipotetiza que si el mayor decremento de los hiperactivos se produce en las TEC o tareas con una tasa muy elevada de targets (y, por tanto, con continuas demandas de memoria de trabajo) se demostrará que el déficit de los niños hiperactivos radica en el sistema de *arousal* (procesamiento inicial de la información). Por contra, si el decremento acontece en tareas con una tasa mucho más reducida de targets (más bien TVI), con poca carga de memoria de trabajo pero con necesidad de mantener a punto el sistema de respuesta, entonces el déficit estaría en el sistema de activación.

La revisión de Van der Meere (1996) muestra como en el primer tipo de tareas las diferencias entre hiperactivos y normales, en su conjunto, han sido amplias y desfavorables lógicamente a los hiperactivos, sin embargo cuando se ha analizado el nivel de decremento atencional la curva ha sido similar. Por tanto, de acuerdo con el modelo expuesto, no es posible hablar de un déficit de atención sostenida en el sistema *arousal* de los niños hiperactivos. Por contra, en el segundo tipo de tareas el grupo de Van der Meere sí ha encontrado no sólo diferencias en el total de las puntuaciones, sino un decremento más rápido y mayor en los niños hiperactivos. En conclusión, el déficit de atención sostenida de los niños hiperactivos se centra especialmente en el sistema de activación. Trabajos posteriores también indican que mientras niños con déficit de atención más trastornos de conducta presentan un mejor funcionamiento del sistema de esfuerzo que les permite compensar sus problemas atencionales, los hiperactivos puros parecen no poderse valer de esta posibilidad. Van der Meere (1996, p. 133 y ss.) también señala que se

han encontrado pruebas favorables a su hipótesis con medidas psicofisiológicas como los potenciales evocados.

En cualquier caso estos estudios, y prácticamente todos los que han utilizado pruebas de medida de la atención sostenida, se enfrentan en la práctica con problemas operativos: en el caso anterior todo el peso recae en la función de decremento atencional, pero como hemos visto la función de capacidad también requiere algún tipo de explicación. Normalmente no disponemos de mucha información psicométrica sobre las tareas de atención sostenida: hay pocos datos de normalización de las puntuaciones por niveles de edad e, incluso, hay muchos problemas para decidir qué tipo de puntuación utilizar. En algunos estudios sólo se valen de los tiempos de reacción (en general una medida demasiado simple cuando se trabaja con muestras normales), en otros sólo de los aciertos (¿qué pasa con un niño que tiene un buen nivel de aciertos pero ha cometido gran cantidad de errores de comisión y que, por tanto, ha podido valerse de una estrategia arriesgada o azarosa?) y en otros sólo de las falsas alarmas (¿qué relación mantienen con los aciertos?). Como veremos en el siguiente apartado, la teoría de detección de señales ha venido a paliar parte de la problemática operativa y, por consiguiente, abrir nuevas vías a la investigación aplicada de la atención sostenida, tanto para el desarrollo de modelos explicativos como para la mejora de los procesos de evaluación, diagnóstico y tratamiento de los niños con déficit atencional.

3.3.2 Sistemas de clasificación: la teoría de detección de señales

La teoría de detección de señales (TDS) «*constituye un modelo psicofísico para evaluar la actuación humana en una amplia variedad de campos como la discriminación sensorial, la percepción y la memoria*». (Reales y Ballesteros, 1997, p. 27), y muy especialmente, por lo que a nuestros intereses respecta, para las tareas de evaluación de la atención sostenida. La TDS nació en el campo de la ingeniería, en el contexto de la detección de señales electromagnéticas en presencia de «*ruido*» (esto es, la detección de señales contaminadas por otras no deseadas). Fue propuesta por primera vez por Peterson, Birdsall y Fox en un simpo-

sium sobre teorías de la información organizado en 1954 por el Massachusetts Institut of Technology, y el mismo año Tanner y Swets lo incorporaron a la investigación psicofisiológica.

El diseño más simple de la TDS, conocido como **diseño de intervalo**, es prácticamente coincidente con lo que es la evaluación de la atención sostenida a partir de las TEC o las TVI: el sujeto, como se ve en el siguiente ejemplo en forma de tabla, se expone a una situación de detección de una señal en presencia de señales distractoras o ruido, con cuatro posibles tipos de respuesta.

Tipo de Ensayo	Tipo de Respuesta	
	SÍ	NO
S+R (señal + ruido)	Acierto	Error de Omisión
R (ruido, distractor)	Error de comisión o Falsa Alarma	Rechazo correcto

Las principales diferencias con las TEC y las TVI es que en éstas ante la presencia del target no siempre se da el ruido, sino que se presentan alternativamente los estímulos señal y los distractores. Por otra parte, en las TEC realmente se da una estructura de respuesta similar al diseño de intervalo, mientras en la TVI en lugar de «rechazos correctos» cabe la «no respuesta».

Como explica extensamente la revisión de Tudela (1987), la TDS es una teoría alternativa a las teorías del umbral, desarrolladas en el ámbito de la investigación psicofísica y sensorial, que intentaban dar cuenta de los procesos cognitivos de detección y respuesta ante estímulos. La investigación arranca de los estudios clásicos sobre el tiempo de reacción y llega hasta nuestros días. Sin entrar en detalle en el complejo entramado teórico que el lector puede consultar en la revisión citada, si nos interesa destacar que desde la psicología del procesamiento de la información se ha aplicado la llamada «**teoría de los dos procesos**» para explicar qué hace realmente el sujeto ante una tarea de detección de señales. Por otra parte, como se verá a continuación, el lector no encontrará muchas difi-

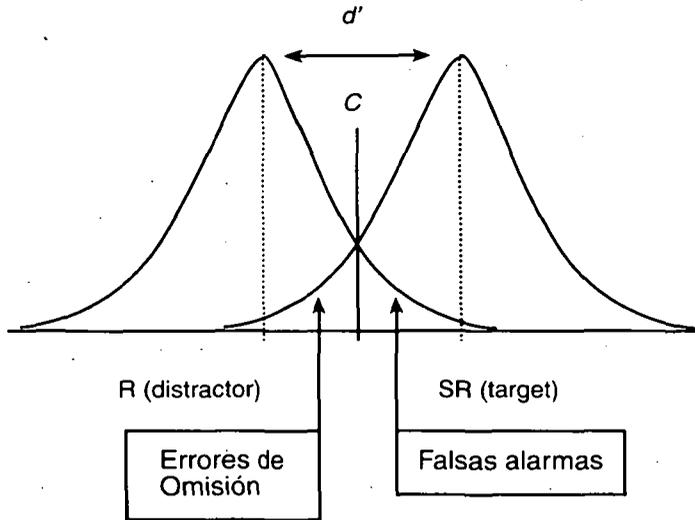
cultades en hacer una relativa equiparación entre cada uno de estos dos procesos y la diferenciación entre los sistemas de arousal y activación que comentamos anteriormente al hilo de los trabajos de Van der Meere.

La teoría de los dos procesos supone que entre la presentación del estímulo (E) y la respuesta del sujeto (R) no medía un único proceso de detección/decisión sino que al menos son dos: tras la presentación del E el sujeto pone en marcha un *proceso sensorial* que le lleva a la generación de un output (X) o «*hipótesis de respuesta*». Sobre el X el sujeto aplica un *proceso decisional* que es el que finalmente le llevará a emitir la respuesta observable. De modo más específico, dentro del proceso sensorial se dan a su vez dos operaciones: la de *codificación* del estímulo presentado y la de búsqueda del estímulo señal o target en la memoria de trabajo. Dentro del proceso decisional se dan también dos operaciones: la de *toma de decisión* (¿es o no es el estímulo que he codificado el target?) y la de *organización de respuesta* (tras la decisión debo dar la respuesta motora adecuada). Como decíamos, es posible hacer una identificación, por una parte, entre el proceso sensorial y el sistema de *arousal* (primera fase de procesamiento de la información, de capacidad de detección y discriminación) y, por otra parte, entre el proceso decisional y el sistema de activación (segunda fase de procesamiento, de tipo motor, de ejecución de la respuesta).

Para realizar sus cálculos, la TDS parte del supuesto inicial de que el sistema no funciona por umbrales, sino que siempre manifiesta algún tipo de activación. Por tanto, la forma que adoptará la distribución de las probabilidades de los valores de sensación, al contrario de lo que pasa en las teorías de los umbrales, no tiene una especial relevancia: de hecho en función del tipo de tarea y del propio sujeto esta distribución puede adoptar formas muy diferentes. En cualquier caso, la figura 1 muestra la distribución más habitual, acompañada de los dos índices básicos que el investigador debe evaluar: la sensibilidad a la tarea y el criterio de decisión del sujeto (también llamado sesgo de respuesta o razón de verosimilitud).

En la figura se aprecian las distribuciones de los ensayos con presencia del target (SR) y con presencia de los estímulos distractores o ruido

Fig. 1. Distribuciones de ensayos en presencia del target (SR) y del distractor (R) en un tarea de detección de señales y cálculo gráfico del índice de sensibilidad (d') y del sesgo de respuesta o criterio (C).



(R). En este caso, por ser el ejemplo prototípico, las distribuciones son muy parecidas. Sobre ellas, justo en el cruce de ambas curvas, se pueden observar a un lado la tasa de errores de omisión (es decir, los targets no detectados) y a otro lado la tasa de falsas alarmas o errores de comisión (es decir, cuando el sujeto ha creído ver equivocadamente el target). También ambas proporciones de errores son aquí muy parecidas pero obviamente en función de las tareas o las diferencias individuales podrían ser muy variables. Las dos medidas de evaluación a las que aludíamos, y que son de naturaleza independiente, son las siguientes:

- **El índice sensibilidad a la tarea (d'):** matemáticamente expresa la distancia entre la distribución de aciertos y la de falsas alarmas. Es decir, es el índice que refleja la capacidad de discriminación del sujeto o, en el tema que nos ocupa, la capacidad de atención sostenida del niño.

- **El criterio de decisión (C) o razón de verosimilitud (Beta):** matemáticamente expresa la preponderancia de la proporción de errores de omisión o la de comisión. Es decir, determina el posible sesgo de respuesta del sujeto en referencia a mostrarse más «conservador» (la proporción de comisiones es menor que la de omisiones) o más «arriesgado» (la proporción de comisiones es mayor que la de omisiones).

Existen diversas propuestas de cálculo matemático de estos dos índices, especialmente en el caso del sesgo de respuesta. Reales y Ballesteros (1997) los revisan e incluso realizan una propuesta propia de cálculo a través de un programa informático muy interesante. En el caso de d' la cuestión está más clara: el procedimiento habitual consiste en transformar a puntuaciones Z las proporciones de aciertos y falsas alarmas del sujeto, según las tablas de la curva normal, y proceder a su resta.

$$d' = \text{Puntuación } Z \text{ de aciertos} - \text{Puntuación } Z \text{ de Falsas Alarmas}$$

La mayor capacidad de discriminación (o mayor nivel de atención sostenida) viene dado por puntuación d' elevadas. Por su parte, el cálculo del sesgo de respuesta viene dado por distintos índices. El más conocido es la razón de verosimilitud o Beta. En principio la base de su cálculo está en la división entre la proporción de aciertos y la proporción de falsas alarmas, pero como señalan Reales y Ballesteros (1997), para realizar comparaciones es necesario estandarizar ambas puntuaciones. Ello se consigue también con los valores de la curva normal; en este caso hay que buscar el valor Y (el eje de ordenadas) que corresponde a cada uno de los valores Z (el eje de abscisas) de las proporciones de aciertos y falsas alarmas.

$$\text{Beta} = Yz \text{ de aciertos} / Yz \text{ de Falsas Alarmas}$$

Dado que el resultado final es un cociente, cuanto más cercano a 1 esté su valor menor sesgo de respuesta se ha detectado. Por contra valores muy alejados de 1 indican un sesgo conservador, la proporción de

omisiones es mayor que la de comisiones. A su vez valores muy cercanos al 0 indican un sesgo arriesgado del sujeto, la proporción de comisiones es mayor que la de omisiones. Recordemos que la proporción de omisiones es la complementaria de la de aciertos que aparece en la fórmula.

Aunque, como decíamos, la razón de verosimilitud es el índice más habitual de sesgo de respuesta, tal y como proponen Reales y Ballesteros (1997), el cálculo del criterio de decisión (C) ofrece la misma información y gráficamente, si se trabaja con puntuaciones transformadas en función de la curva normal, parece más claro. La razón es que la distribución por encima y por debajo del 1 no es simétrica. Obviamente en la curva normal Y sólo puede fluctuar entre 0 y 1, mientras por encima del 1 se pueden llegar a valores por encima de 3 y hasta cercanos a 4, aunque sean poco probables. Así pues, la distancia entre un sujeto con una Beta de 0,5 y otro de 1,5 en absoluto es comparable. En cambio C, aunque matemáticamente haya que añadirle un valor constante, permite perfectamente este tipo de comparaciones. Su cálculo es el siguiente:

$$C = - 0,5 \times (Z \text{ de aciertos} + Z \text{ de Falsas Alarmas})$$

El valor constante es el de -0,5. Éste se multiplica a la suma de las puntuaciones Z de aciertos y falsas alarmas y nos encontramos así con una distribución cuyo punto crítico es el cero; efectivamente cuando las proporciones de errores de omisión y comisión son idénticas (no hay, pues, ningún sesgo de respuesta) el valor de C es cero. Por contra, si la de omisiones es superior el valor de C tiende a ser positivo (sesgo de respuesta conservador) y si es superior la de comisiones C tiende a ser negativo (sesgo de respuesta arriesgado). Por supuesto un sesgo de -0,5 y otro de 0,5 ahora sí expresan igual magnitud, respectivamente, de sesgo arriesgado y conservador.

Aunque sólo sea a título de ejemplo didáctico, con valores Z aproximados, veamos dos casos con sus respectivos cálculos. En primer lugar imaginemos el caso de un sujeto que ha obtenido una proporción del

90% de aciertos (es decir, 10% de omisiones) y del 10% de falsas alarmas. La Z de aciertos aproximada que le corresponde es de 1,282, mientras la Z de falsas alarmas es de -1,282, es decir, son la misma sólo que una está muy por encima de la media normalizada y la otra muy por debajo. Al restar ambas puntuaciones obtenemos una $d' = 2,564$ (alta sensibilidad del sujeto). En cuanto al sesgo de respuesta resulta inexistente: tanto si calculamos Beta como C . El primero tiene el valor 1 resultante de dividir Yz aciertos de Yz de falsas alarmas (ambos con el valor aproximado de 0,176), mientras el segundo tiene el valor 0 resultante de multiplicar la constante -0,5 por la suma de las puntuaciones Z , que en este caso da 0.

En segundo lugar imaginemos otro sujeto con una proporción de aciertos también del 90% ($Z = 1,282$), pero con una de falsas alarmas del 80% ($Z = 0,842$). Así las cosas d' valdría 0,44: el sujeto demuestra ser poco sensible sobre la tarea. Por su parte, Beta valdría 0,63, un sesgo claro hacia el riesgo, que también refleja perfectamente C con un valor de -1,06. Como vemos Beta y C son distintos procedimientos para calcular el mismo concepto, y aunque sea más popular el primero, nosotros, como proponen Reales y Ballesteros (1997), entedemos que el segundo da una información visual más fácil de interpretar.

El problema de los cálculos basados en la curva normal, y del programa que proponen Reales y Ballesteros (1997), es, en primer lugar, que sólo están pensados para adultos y, en segundo lugar, que no sabemos cuál es la puntuación que nos permite hablar ya de sensibilidad o sesgo significativo (no sólo matemáticamente sino de forma aplicada). Al trabajar con muestras de niños de diferente edad, por ejemplo entre 6 y 10 años, si no se tiene en cuenta que su nivel medio de aciertos y falsas alarmas es muy diferente ocurrirá que comparativamente todos los niños de 6 años aparecerán con una muy débil sensibilidad estimular (o si se prefiere, baja capacidad atencional), cuando lo que realmente nos interesa es establecer comparaciones dentro de cada edad y entre edades, pero manteniendo la magnitud proporcional de sus diferencias evolutivas. Para ello nos vendría muy bien disponer de datos normativos y, por ende, de los niveles medidos de proporción de aciertos y falsas

alarmas que cabría esperar en una determinada tarea de vigilancia o de detección de señales pensada para evaluar niños. Ello nos permitiría ajustar los cálculos de d' y C a cada grupo de edad. De este modo, al menos en lo que a la psicología aplicada se refiere, sería mucho más fácil diagnosticar la presencia de un déficit en la capacidad de atención de un niño y/o estudiar el efecto de este mecanismo sobre las variables de aprendizaje y comportamiento que deseamos.

3.4 El trastorno por déficit de atención con y sin hiperactividad

El autor del presente trabajo ha revisado anteriormente las características de los niños con trastornos por déficit de atención con y sin hiperactividad (respectivamente TDA+H y TDA-H), por ejemplo en Servera (1999) o en Servera y cols. (2001). Pero al igual que hemos hecho con el niño de estilo impulsivo, conviene recordar qué caracteriza a un niño con «*déficit de atención*».

En principio hay que recordar que las TEC y las TVI sólo son una parte del diagnóstico de este déficit y los trastornos de hiperactividad en general. Es necesario una batería de evaluación más amplia, de carácter neuropsicológico, junto con escalas de padres, maestros y registros de observación para determinar la presencia de estos problemas. El TDAH puede definirse como *un trastorno del desarrollo caracterizado por niveles de inatención, sobreactividad e impulsividad inapropiados desde el punto de vista evolutivo. Estos síntomas a menudo se inician en la primera infancia, son de naturaleza relativamente crónica y no pueden atribuirse a alteraciones neurológicas, sensoriales, del lenguaje o motoras graves, a retraso mental o a trastornos emocionales severos. Estas dificultades se asocian normalmente a déficits en las «conductas gobernadas por reglas» y a un determinado patrón de rendimiento»* (Barkley 1990, p. 47). A la hora de hacer el diagnóstico, aunque quedan muchos aspectos polémicos por investigar, hay un cierto acuerdo en mantener la diferenciación entre el TDA+H y el TDA-H, pero ¿qué diferencias hay entre ellos?

En primer lugar los niños con TDA+H presentan una sobreactividad motora o, si se prefiere, simplemente una actividad motora desadaptativa muy característica. Realmente les cuesta controlar su conducta motora especialmente en situaciones que exigen esfuerzo cognitivo y atención. Pero, además, en su mayoría también presentan conductas disruptivas e incluso antisociales, claramente implicados en sus problemas académicos y de adaptación familiar y escolar. En segundo, lugar, por lo que respecta a la problemática en las disfunciones atencionales, los clínicos han tenido que acercarse a la investigación básica. Al menos dos tipos de atención, la selectiva y la sostenida, cada vez parecen tener más implicaciones en el proceso de evaluación y tratamiento. En general, y aún siendo un tema en estudio, los niños con un TDA+H demuestran que la mayor parte de sus problemas atencionales se centran en la atención sostenida, mientras los niños con un TDA tienen más problemas en la selectiva. En palabras de Rusell Barkley:

«los niños con TDA-H pueden tener más problemas con la atención focalizada o velocidad del procesamiento de la información (análisis de los inputs y recuperación de la información almacenada). Por su parte, los niños con TDA+H pueden tener más dificultades con la atención sostenida y el control de la impulsividad, y con los parámetros motivacionales implicados en la tarea» (Barkley, 1990, p. 90).

Hasta ahora los datos, especialmente los referentes a los niños con TDA+H tienden a confirmar esta hipótesis. Por tanto, a pesar de que muchos padres y educadores describen como «*distráctiles y con poca capacidad de concentración*» a muchos niños hiperactivos, cuando se ha analizado su rendimiento sobre tareas de laboratorio centradas en la atención selectiva y la discriminación visual se ha observado que en muchos casos no llegan a diferenciarse significativamente de los niños normales (Douglas y Peters, 1979; Taylor, 1986; 1994). En cambio, y de modo mucho más consistente, estos niños presentan niveles de rendimiento significativamente inferiores en tareas de rendimiento continuo y de vigilancia, propias del proceso de atención sostenida (Barkley, 1990; 1997; Moreno, 1995; Taylor, 1986; 1994).

En cuanto a los niños con TDA-H, es decir los que tienen un déficit de atención propiamente dicho, constituyen un grupo mucho menos estudiado. Como señala Barkley (1990) su presencia ha enmascarado tanto los estudios con niños hiperactivos como los estudios con niños con dificultades de aprendizaje. En el primer caso, es evidente que muchas de las contradicciones en las conclusiones experimentales sobre la hiperactividad podrían provenir de la presencia más o menos mayor de TDA-H en las muestras investigadas. Bradley, Grodzinsky y DuPaul (1992) encuentran evidencias claras en favor de características propias de estos niños: menos del 20% presentarían problemas conductuales más o menos graves, prácticamente no tiene problemas de sobreactividad motora, ni se diferencian significativamente de los niños normales en tareas de atención sostenida, o de funcionamiento del lóbulo frontal en general. En cambio, sí se caracterizan por problemas de rendimiento cognitivo: tanto en tareas académicas, como de laboratorio, especialmente cuando se requiere un esfuerzo de concentración y focalización atencional.

Por lo que respecta a la relación entre el déficit atencional y las dificultades de aprendizaje, también existen bastantes problemas. Como afirman López-Soler y García-Sevilla (1997), entre el 40-50% de las dificultades de aprendizaje se asocian a problemas atencionales, hasta el punto que algunos autores consideran los trastornos hiperactivos como sinónimos del término «*dificultad de aprendizaje*» o viceversa. Desde un punto de vista genérico, los niños hiperactivos coinciden con muestras de niños con dificultades de aprendizaje específicas en que, en general, sus puntuaciones sobre el CI y otras medidas cognitivas globales entran en los intervalos de normalidad, mientras sus puntuaciones en pruebas de rendimiento académicas (normalizadas y basadas en el currículum) suelen estar 1,5 desviaciones estándar por debajo de la media. También hay coincidencia en que los dos grupos de niños presentan con frecuencia conductas de falta de atención muy evidentes. Sin embargo, no parece del todo adecuado equiparar hiperactividad y dificultad de aprendizaje sin realizar algunas matizaciones.

En el TDA+H no tiene porqué estar presente una afectación específica del área perceptiva, el lenguaje o la psicomotricidad (las más típicas

en las dificultades de aprendizaje) (López-Soler y García-Sevilla, 1997). Además, se supone que las pruebas atencionales que mejor les caracterizan son las de ejecución continua y vigilancia (atención sostenida), mientras en el caso de las dificultades de aprendizaje destacan más las medidas de distraibilidad y de focalización atencional (atención selectiva). Evidentemente, en este último aspecto sí es posible una cierta equiparación entre dificultad de aprendizaje y TDA puro, sin embargo hay que tener presente que en este último caso normalmente no puede determinarse un déficit cognitivo específico responsable de sus problemas académicos. La función atencional, como hemos visto, no se considera un proceso cognitivo, sino la actividad básica que permite la puesta en marcha de todo el procesador central de información (Ruiz-Vargas y Botella, 1987). Por tanto, su disfunción tiene un carácter inespecífico y global sobre la ejecución y el rendimiento.

Por tanto, y a modo de resumen, podemos decir que el diagnóstico de trastorno por déficit de atención hace referencia a un niño con una capacidad cognitiva general dentro de la normalidad, pero que tiene problemas para detectar los estímulos relevantes, comete errores por descuido, se distrae fácilmente, parece no escuchar, tiene dificultades para seguir instrucciones y tiene problemas para mantener la concentración en tareas repetitivas de larga duración (aunque esta última característica suele ser indicativa de la presencia de un TDAH combinado). En el TDA-H no tiene porqué haber conductas problema importantes ni en casa ni en la escuela, pero desde un punto de vista estricto, para su diagnóstico el DSM-IV exige actualmente que las disfunciones atencionales sean percibibles en ambos ambientes (o más) y que supongan un deterioro clínicamente significativo de la actividad social, académica o laboral (APA, 1995). De todos modos, sea cual sea la naturaleza de la disfunción atencional, tanto si pertenece a un déficit puro, como si se incluye en un TDAH combinado, la evaluación y la mejora de la capacidad de mantenimiento del foco atencional parece ser un pilar básico del trabajo clínico y educativo. Como ya hemos intentado justificar, a pesar de que se puede asumir la relativa independencia de los distintos tipos de atención y sus funciones, el mantenimiento in-

corpora el mínimo de esfuerzo y recursos que exigen todos los restantes mecanismos. Por otra parte, en el trabajo escolar de cada día la habilidad para mantener la atención en tareas repetitivas, automatizables hasta cierto punto, pero que requieren no cometer errores es tan habitual como en nuestra vida diaria.

Para finalizar nos gustaría hacer hincapié en que todo lo dicho anteriormente para la disfunción atencional es bastante equiparable a lo comentado ya al hablar del estilo impulsivo. En sentido estricto ambas dimensiones pueden ser consideradas un tipo de dificultad de aprendizaje, pero dada su implicación básica para el funcionamiento de todos los demás procesos de solución de problemas y mecanismos cognitivos asociados, hay que entender que su influencia es mucho más vasta y compleja que la que afecta a las dificultades de aprendizaje habituales. Es más, la presencia de un déficit de atención o un problema de impulsividad en el niño puede ir más allá de lo que es su repercusión inmediata (a veces incluso puede pasar relativamente desapercibida para padres y educadores sin la presencia de conductas problema graves que, como hemos visto, en muchos casos no se dan) y llegar a afectar solapadamente a todo lo que es el desarrollo evolutivo de sus sistemas de procesamiento de la información y funcionamiento intelectual.

PARTE EXPERIMENTAL: EL DESARROLLO DE MEDIDAS OBJETIVAS PARA LA EVALUACIÓN DE LA IMPULSIVIDAD COGNITIVA Y LA ATENCIÓN SOSTENIDA

4. LOS OBJETIVOS DEL TRABAJO

En la primera parte de este trabajo hemos revisado la conceptualización y los problemas de evaluación de dos de las dimensiones más implicadas en los problemas escolares del niño: el estilo reflexividad-impulsividad y la atención sostenida. La conclusión que creemos importante destacar, en primer lugar, es que se trata de dos dimensiones complejas, enraizadas en constructos multicomponentes, que han generado gran cantidad de investigación pero que no han acabado de concretarse en aspectos aplicados que puedan ser fácilmente utilizados por los educadores. En segundo lugar, pero relacionado con este primer aspecto, también entendemos que gran parte de esta situación deriva de una confusión en la propia definición y evaluación de las dos dimensiones. Para empezar el sentido común relaciona la impulsividad del niño con la inatención, pero es que realmente ambas disfunciones son aplicadas a la inmensa mayoría de niños con problemas de aprendizaje y/o comportamentales. Si todo niño con un problema en el aula tiene, por definición un déficit de atención o un estilo impulsivo, estas dimensiones pierden gran parte de su sentido.

En nuestra revisión hemos intentado dejar claro que existen distintos tipos de impulsividad y de problemas de inatención, que cada uno puede tener implicaciones diferentes (aunque se pueden dar conjuntamente) y que, con datos en la mano, no es fácil concluir el tipo de interacción mantienen sobre el aprendizaje y el comportamiento del niño. Para em-

pezar a abordar este problema, al menos desde nuestra óptica, se hace necesario definir pruebas objetivas para evaluar cada tipo de impulsividad y de inatención. Estas pruebas, llamadas de laboratorio, tienen limitaciones evidentes, al derivar de tareas *artificiales* muy poco habituales en la vida cotidiana del niño, pero también ventajas importantes cuando necesariamente se requiere la evaluación de una base de actuación cognitiva, difícil de concretar a través de registros de observación, informes o escalas de evaluación para adultos. Por ejemplo, en nuestro caso que nos hemos decantado por el estilo de procesamiento de información reflexividad-impulsividad y por el mecanismo de la atención sostenida hemos encontrado suficientes evidencias que a través de la conducta manifiesta del niño no se puede inferir ni descartar la presencia de estos tipos de disfunción, cuando, por otra parte, su presencia sí parece tener implicaciones directas al menos en dificultades de aprendizaje, y en cierta medida comportamentales.

Una vez se asume la necesidad de utilizar pruebas objetivas para la evaluación de la impulsividad y la inatención la cuestión estriba, como creemos haber expuesto justificadamente, en los problemas metodológicos para su elaboración, su validación y los criterios de decisión para clasificar a los sujetos. En el caso de la elaboración las bases que las sustentan están muy claras desde hace tiempo, aunque en general o bien se trata de pruebas con problemas psicométricos (caso de la reflexividad-impulsividad), o bien pruebas con dificultad para pasar de la investigación experimental a la aplicada (más habitual en la evaluación de la atención). Los problemas de validación y los criterios de clasificación de los sujetos derivan de una raíz común: las pruebas de laboratorio necesitan criterios fiables para diferenciar el funcionamiento de los niños y, a su vez, validarse a través de la constatación de diferencias entre ellos en aspectos de la vida académica y social. Todas estas cuestiones, en mayor o menor medida, como hemos visto en las páginas precedentes, presentan cabos sueltos.

Ante esta situación nosotros nos planteamos una serie de objetivos generales y específicos que son los que han impulsado nuestro trabajo de investigación. Los objetivos generales pueden resumirse de la siguiente manera:

1. Elaborar una prueba objetiva para la evaluación del estilo cognitivo reflexividad-impulsividad que, aunque derive de las versiones existentes hasta el momento, sea aplicable por ordenador y permita un cálculo más sencillo de sus puntuaciones directas y sus índices combinados.
2. Elaborar una prueba objetiva (también por ordenador) para la evaluación de los problemas de atención sostenida que, aunque derive de los paradigmas experimentales de las tareas de vigilancia y ejecución continua, pueda aplicarse habitualmente en los centros escolares. En este caso los índices combinados derivarán de la aplicación de la teoría de detección de señales.
3. Analizar las implicaciones de las medidas y los sistemas de clasificación de la reflexividad-impulsividad sobre el rendimiento académico y el comportamiento del niño.
4. Analizar las implicaciones de las medidas y los sistemas de clasificación de la atención sostenida sobre el rendimiento académico y el comportamiento del niño.
5. Analizar las relaciones entre la reflexividad-impulsividad y la atención sostenida.

Los objetivos específicos son los siguientes:

1. Comprobar las diferencias por edad y por sexos de las distintas medidas e índices de la reflexividad-impulsividad y la atención sostenida.
2. Analizar las diferencias entre los distintos sistemas de clasificación de sujetos impulsivos y reflexivos sobre las variables de rendimiento y comportamiento del niño. Estos sistemas son el de cuadrantes, el integrado y el IMAT, éste último propuesto por nosotros.
3. Analizar la aportación de los índices de sensibilidad y sesgo de respuesta (derivados de la teoría de detección de señales) frente a las medidas directas de la prueba de atención sostenida sobre variables de rendimiento y comportamiento del niño.

4. Analizar las posibilidades de clasificación de sujetos en baja y alta capacidad atencional a través de los índices de sensibilidad y sesgo de respuesta.

A continuación se detallan los aspectos metodológicos de nuestro trabajo, con el análisis de los resultados obtenidos, y terminaremos con un análisis de las conclusiones que hemos podido extraer.

5. METODOLOGÍA

5.1 Sujetos

Para obtener la muestra que ha servido de base al presente estudio se evaluaron inicialmente a 978 niños de primero a cuarto de educación primaria procedentes de seis colegios públicos y concertados de Mallorca. Todos los niños fueron evaluados de manera individual en dos pruebas por ordenador: la tarea de emparejamiento de figuras (TEF-IMAT), para la medida de la dimensión reflexividad-impulsividad, y la tarea de vigilancia (TVI-IMAT). Además se recogieron los datos de una escala de elaboración propia que cumplimentaron los tutores de cada niño: la Escala de Comportamiento y Rendimiento en el Aula (ECRA). Con ayuda de los evaluadores y de los tutores se realizó una primera criba de la muestra inicial por la que sólo se mantuvieron en el estudio los niños que cumplían las siguientes condiciones:

- No haber presentado ningún problema comportamental grave en la realización de las pruebas por ordenador.
- No estar en la categoría de «necesidad educativa especial», no pertenecer a programas de integración, ni padecer cualquier tipo de alteración que pudiera condicionar los resultados de las pruebas de evaluación.
- No presentar cualquier otra característica que quedase reflejada en la escala para maestros (ECRA) o en los registros de nuestros evaluadores que se considerase influyente en los resultados finales de la prueba.

A partir de los datos codificados, y dado que nuestro interés era poder extraer el máximo de conclusiones generalizables del estudio se procedió a una «depuración» de la muestra eliminando a los sujetos *outliers* o extremos en las siguientes variables (dentro de cada curso): edad, errores y latencias de la TEF, aciertos, errores de comisión y tiempo de reacción en la TVI. De este modo, sin duda, en muchos análisis estadísticos se observará cierta variabilidad del número de sujetos que componen los grupos, pero en todos los casos cabe hablar, en nuestra opinión, de muestras razonablemente grandes. A modo de orientación en la tabla 5 se presenta la distribución de sujetos cuando se trabaja específicamente con las medidas de la prueba TEF o de la TVI.

Tabla 5. Distribución de la muestra para las variables de la TEF y la TVI

Curso/Grupo	Niños	Niñas	N	Media Edad	Desv. Est.	Mínimo	Máximo
TEF							
1°	162	98	260	7,11	0,30	6,10	7,48
2°	113	93	206	8,00	0,29	7,50	8,50
3°	75	55	130	9,04	0,32	8,50	9,48
4°	92	82	174	10,01	0,30	9,51	11,23
total	442	328	770	8,33	1,16	6,10	11,23
TVI							
1°	116	77	193	7,10	0,30	6,10	7,48
2°	84	62	146	8,00	0,28	7,50	8,49
3°	59	43	102	9,04	0,30	8,51	9,48
4°	76	67	143	10,02	0,31	9,51	11,23
total	335	249	584	8,38	1,18	6,10	11,23

Es importante, pues, hacer énfasis, además de que en casi ninguna variable se encuentran implicados todos los sujetos (ya sea por quedar fuera a causa de una puntuación extrema, o porque falte esa puntuación), que la variable «*curso*» en realidad hace referencia a «*grupos de edad*» que, como se ha podido observar, intenta discriminar al máximo a cada uno de los subgrupos implicando el mínimo de solapamiento posible.

5.2 Las pruebas de evaluación

5.2.1 La Tarea de Emparejamiento de Figuras (TEF-IMAT)

La TEF-IMAT (a partir de ahora sólo TEF) es una versión totalmente renovada del clásico MFF que como principales características presenta el hecho de aplicarse por ordenador y de constar de 16 ítems nuevos. La aplicación por ordenador permite una mayor precisión en la medida de la latencia de respuesta, una simplificación en la forma de aplicación y, sobre todo, un sistema de corrección automatizado. Con la TEF se puede realizar la clasificación de sujetos por el procedimiento de cuadrantes, por el sistema integrado de Salkind y Wright (1977) o por el sistema IMAT, que es el que proponemos nosotros, y que explicaremos más adelante.

La razón de funcionar con 16 ítems se encuentra en los estudios pilotos que se realizaron hasta llegar a la versión definitiva. Con la elaboración inicial de 60 ítems, la prueba fue aplicándose a amplias muestras de sujetos con la intención de seleccionar aquellos con mayor poder de discriminación. Finalmente, de una primera versión elaborada de la TEF de 40 ítems se obtuvieron datos de 974 niños de edades comprendidas entre 6 y 11 años y seleccionaron los 16 ítems actuales, que cumplieron los siguientes criterios: (a) presentaban una correlación con el total de la prueba igual o superior al .35, y (b) estaban entre los márgenes de ± 1 desviación estándar en relación a la media total de errores (es decir, se eliminaron los ítems más fáciles y los más difíciles). También se

comprobó que los índices de fiabilidad y la correlación errores-latencias no sufrían variaciones importantes con la reducción de ítems.

Los 16 ítems de la TEF presentan idéntica estructura. El niño observa en la parte superior de la pantalla del ordenador una figura modelo, y debajo de una línea de separación, otras seis figuras muy similares de entre las cuáles sólo una es idéntica al modelo. Su tarea consiste en mover el cursor de la pantalla con el «ratón» del ordenador y hacer «clic» en la figura que, según él, es la correcta. El ordenador registra de manera interna para cada ítem el número total de errores y la latencia de la primera respuesta (en milisegundos). Si la primera respuesta es acertada, aparece una cara sonriente en la pantalla, y automáticamente se avanza hacia el siguiente ítem. Si no es así, aparece una cara triste en la pantalla y vuelve al mismo ítem; esta vez cuando el cursor pase por encima de la alternativa aparecerá la palabra «NO», de manera que no se permite que cometa el mismo error. Antes de los ítems de evaluación aparecen cuatro de ejemplo en los que el evaluador se cerciora de que el niño ha entendido las instrucciones y de que domina correctamente el ratón. Normalmente se permiten tres pruebas sobre estos ítems ejemplo, pero nada impide que se puedan realizar más (u otro tipo de ejercicios) hasta que el niño muestre autonomía en el uso del ratón. En aproximadamente las dos mil evaluaciones realizadas hasta llegar a la versión definitiva sólo en casos muy excepcionales, incluyendo a los niños entre 6 y 7 años, hemos tenido que suspender una aplicación porque el niño presentase un uso poco habilidoso del ratón. Hemos realizado controles para comprobar si la experiencia previa en el manejo de ordenadores podía condicionar el tiempo de respuesta y han sido negativos.

Al finalizar la prueba el evaluador tiene acceso a las puntuaciones directas de errores y latencias del sujeto que permiten el cálculo de los índices de estilo (PI) y eficacia (PE).

5.2.2 La Tarea de vigilancia (TVI-IMAT)

La TVI-IMAT (a partir de ahora TVI) es una versión propia de las tareas de vigilancia, a veces denominadas erróneamente «tests de ejecu-

ción continua» o CPT (*continuous performance test*), utilizadas para la evaluación de la capacidad de atención sostenida. En esencia estas tareas obligan al sujeto mantener un cierto nivel de activación y de respuesta sobre tareas repetitivas de larga duración. Como vimos en su momento, en sentido estricto para que una de estas tareas mida realmente atención sostenida debe durar lo suficiente para que todos los sujetos experimenten «*cansancio*», esto es, disminución de rendimiento en mayor o menor grado. Esto en adultos suele suponer márgenes de trabajo constante mínimo entre 20 y 30 minutos. En niños los estudios son menores y no existen pautas claras sobre el tipo de tarea y la duración. Además el margen de variabilidad de su potencial de atención sostenida es mucho mayor que en adultos lo que complica esa decisión. Tras dos años de estudios pilotos, en donde fuimos variando todos los parámetros implicados en las tareas de vigilancia buscando una prueba que fuese aplicable entre 6-7 años y 10-11 años, decidimos elaborar la versión actual de la TVI. Esta versión funciona en el ordenador proyectando la pantalla números del 0 al 9, de uno en uno, con los siguientes parámetros:

- (a) target o estímulo objetivo = doble, presencia de un 3 precedido por un 6, (6-3)
- (b) número de estímulos que se presentan en pantalla = 600 (300 parejas)
- (c) duración del estímulo en pantalla = 250 milisegundos
- (d) intervalo interestimular = 500 milisegundos
- (e) Porcentaje de presentaciones del target = 30%
- (f) duración total de la prueba = 7 minutos y 30 segundos.

La labor del niño es apretar la barra de espacio del ordenador cada vez que en la pantalla ha visto un 3 precedido por un 6. El parámetro más conflictivo sin duda es la duración de la prueba, demasiado corta para los teóricos del tema. Ciertamente no podemos demostrar que con esa duración se produzca el efecto de «*menoscabo*

atencional» que se exige para considerar a la tarea una verdadera tarea de vigilancia y, por tanto, de medida de la capacidad de atención sostenida. En nuestro descargo hay que decir que al tratarse de una investigación aplicada y desarrollada en el ámbito escolar, no en el laboratorio, ya sabíamos de entrada que una prueba con una duración de unos 30 minutos, desarrollada en los centros escolares, y con muestras de sujeto cercanas a los 1.000 niños sería casi imposible de aplicar. Además al pretender abarcar edades entre primero y cuarto de primaria probablemente la duración de la prueba hubiera tenido que ajustarse para cada curso, puesto que los efectos madurativos sobre el mecanismo atencional son drásticos, lo cual aún era más complicado para nuestro trabajo. Es por ello que desde el principio realizamos nuestros estudios pilotos no tanto con la idea de que la prueba garantizara el menoscabo atencional, sino la capacidad para discriminar sujetos: aquéllos con mayor y menor capacidad atencional (sea o no exactamente considerada como «*sostenida*») sobre una medida de laboratorio. No obstante, y aún sin disponer de datos definitivos (estamos trabajando sobre ellos actualmente), la observación comportamental de los niños puso de manifiesto que la mayoría experimentaba la sensación de cansancio y aburrimiento habitual en este tipo de tareas (por supuesto mucho más los más pequeños que los mayores). Al finalizar la aplicación de la prueba el ordenador registraba las siguientes puntuaciones directas:

- Aciertos: el número de targets detectados por el sujeto sobre un máximo de 90.
- Omisiones: los targets que han pasado por alto al sujeto, es decir, 90 n° aciertos.
- Comisiones: los errores que derivan de responder sin la presencia del target. Es decir, las falsas alarmas. En principio se presentaban 210 parejas de estímulos que no eran targets.
- Tiempo de reacción en los aciertos: en milisegundos.
- Tiempo de reacción en las comisiones: en milisegundos

5.2.3 *La Escala del Comportamiento y el Rendimiento en el Aula (ECRA).*

La ECRA es una escala de elaboración propia pero con una estructura e ítems similares a otras conocidas escalas de valoración comportamental de los niños en el aula. Veamos, a continuación, su estructura tal y como se presentaba a los tutores de los niños para su cumplimentación:

Los ítems de la ECRA se han escogido de las siguientes escalas: en primer lugar, el CAP (*Child Attentional Profile*), el perfil de problemas atencionales derivado del conocido *Child Behavior Checklist* del grupo de Achenbach y Edelbrock (1983). En segundo lugar, el SNAP (véase Kirby, 1986), un listado de síntomas de hiperactividad basado en el DSM-III pero la mayoría presentes en todos los sistemas clasificatorios de este trastorno. Y, finalmente, los criterios específicos del DSM-IV para el trastorno por hiperactividad con déficit de atención. Más específicamente podemos comentar que el contenido de todos los ítems del CAP está presente en la subescala de problemas de inatención de la ECRA, en el caso del SNAP se han seleccionado y adaptado (no traducido) los más comunes y propios de la hiperactividad y, finalmente, en el caso del DSM-IV se hallan presentes sus 9 criterios de inatención, también sus tres de impulsividad y, en cambio, en el caso de sobreactividad motora sólo aparecen tres. No aparecen: *el niño salta o corre en exceso; tiene problemas para jugar sólo; y parece como si tuviera un «motor»*. La razón para ello es que éstos tienen ya un carácter más estrictamente clínico y la ECRA pretendía solamente una exploración general del comportamiento y el rendimiento del niño. Por esta razón, se añadieron los ítems de la subescala de problemas de trabajo escolar y rendimiento académico.

Aunque es una escala propia, y por tanto sin datos específicos de validación, el hecho de basarse en ítems de otras escalas ampliamente utilizadas y los datos que a continuación ofreceremos da cierta entidad a su utilización. En primer lugar hay que destacar que, en general, existe una elevada relación entre las tres escalas de comportamiento y la de problemas de trabajo escolar: los coeficientes de correlación fluctúan

entre .61 (entre impulsividad y problemas de trabajo escolar) y .87 (entre inatención y problemas de trabajo escolar). Por su parte, la subescala de rendimiento correlaciona alrededor del -.65 con las de inatención y problemas de trabajo escolar, pero en cambio ofrece valores más bajos con impulsividad (-.28) y problemas de interacción social (-.39). En segundo lugar, aunque no hemos explorado en profundidad la composición factorial de nuestra escala, podemos comentar que del análisis por componentes principales de los 23 ítems emergieron tres grandes factores. En el primer factor, que podemos denominar de «*comportamiento y rendimiento general*» y que explica el 56% de la variancia, saturan todos los ítems, con valores que oscilan desde los -.61/- .65 (de los ítems de rendimiento) al .60/.84 (de los ítems de comportamiento). El segundo factor es el de «*impulsividad*» (explica el 13% de la variancia) y en él destacan las correlaciones de todos los ítems de la subescala de impulsividad más el primero de la de problemas de interacción social (valores entre .44 y .54). Finalmente el tercer factor es el de «*rendimiento académico*» (explica el 4% de la variancia) en el que destacan las correlaciones entre .37 y .44 de los tres ítems que valoran el rendimiento.

En definitiva, aunque no hemos encontrado apoyo factorial completo para el mantenimiento de todas las subescalas propuestas, la investigación previa de los instrumentos que nos han servido de base para nuestra escala y los objetivos que pretendíamos en el presente trabajo nos han animado a mantener la estructura. Además la consistencia interna de todas las subescalas ha sido muy elevada, con una amplia correlación de cada ítem con el total de su subescala de referencia: la subescala de inatención presenta una consistencia interna por Alfa de Cronbach de .94. La subescala de impulsividad la presenta de .92. La subescala de problemas de interacción social de .91. La subescala de problemas de trabajo escolar de .89. Y, finalmente, la subescala de rendimiento académico de .93. Por todo ello consideramos que la ECRA es un buen referente de la opinión del maestro respecto del comportamiento y el rendimiento del niño en el aula, y puede servir de criterio para analizar el grado de validez de las pruebas TEF y TVI en el ámbito educativo.

ESCALA DEL COMPORTAMIENTO Y EL RENDIMIENTO EN EL AULA (E.C.R.A.)

Este cuestionario pretende obtener información sobre posibles problemas de comportamiento o rendimiento de sus alumnos. Si no se indica lo contrario, usted deberá valorar todas las afirmaciones en función de la siguiente escala:

- A** Casi nunca, **B** Pocas veces, **C** A veces,
D Bastantes veces, **E** Muchas veces

Cuando aparezca alguna pregunta que no pueda responderse con esta escala, se le indicarán nuevos valores posibles de su respuesta. En cualquier caso no escriba nada aquí: ponga el nombre del alumno/a en la Hoja de Respuestas y marque allí todas sus respuestas. Procure evaluar siempre en función de su propia experiencia en el trato del alumno/a y en referencia a lo que usted ha observado durante el presente año académico.

1. No acaba las cosas (juegos, actividades,..) que el/ella mismo/a empieza.
2. No puede concentrarse en una tarea mucho tiempo.
3. Se pierde en sus pensamientos como si estuviera soñando despierto.
4. Tiene dificultades para seguir instrucciones, parece que no escuchara.
5. Sus trabajos son descuidados y poco pulcros.
6. Se distrae muy fácilmente por cualquier cosa.
7. No hace los deberes que se le asignan.
8. No puede estar sentado, es muy inquieto y se mueve mucho.
9. Es nervioso: siempre manosea algo, hace muecas, mueve extremidades, etc.
10. Es impulsivo y actúa sin pensar.
11. Interviene cuando no es su turno o le cuesta esperar turno.
12. En general habla demasiado.

13. Insulta, amenaza o discute con facilidad.
14. Es rechazado por algunos de sus compañeros.
15. Molesta a sus compañeros e interrumpe sus actividades o la clase.
16. Tiene dificultades para trabajar o participar en grupo.
17. Comete errores «tontos» que es capaz de corregir si se los señalan.
18. Tiene dificultades para organizar su trabajo escolar.
19. Muestra disgusto ante tareas o actividades que implican esfuerzo mental.
20. Pierde u olvida instrucciones, objetos o material importante para su trabajo escolar.
21. ¿Este niño/a presenta algún problema físico, sensorial, psíquico o social (problemas familiares, etc.) que influye muy claramente en su rendimiento académico? (A No, B Físico, C Sensorial, D Psíquico, E Social).

A continuación valore el rendimiento del niño/niña en las siguientes áreas y siempre en función del nivel de la clase (A Muy Bajo, B Bajo, C Normal, D Alto, E Muy Alto)

22. El área de matemáticas, cálculo y resolución de problemas.
23. El área de lectura y comprensión lectora.
24. El área de escritura y expresión escrita.

Nota: el valor numérico de la escala es de A=1, B=2, C=3, D=4, E=5.

Del conjunto de estos ítems se forman las siguientes subescalas:

- Inatención: ítems del 1 al 7.
- Impulsividad: ítems del 8 al 12.
- Problemas de interacción social: ítems del 13 al 16.
- Problemas de trabajo escolar: del 17 al 20
- Rendimiento académico: del 22 al 24.

5.3 Procedimiento

Todos los niños que participan en este estudio fueron evaluados individualmente por los autores con ayuda de alumnos de últimos cursos de carrera del Departamento de Psicología de la UIB. Las ECRA fueron cumplimentadas por los tutores de cada alumno. Naturalmente se respetaron las normas básicas de aplicación de pruebas y a su vez se llevó un registro comportamental de cada niño para poder dar por válido sus resultados.

En *el caso de la TEF* las instrucciones de aplicación fueron las siguientes: al entrar en la primera pantalla de práctica el evaluador explicaba al niño que *«ahora aparecerá una pantalla que en la parte superior tiene un dibujo-modelo, y en la inferior otros seis muy parecidos aunque sólo uno es idéntico al modelo. Tendrás que moverte con el ratón del ordenador para ir activando las alternativas y cuando creas que la has encontrado aprieta este botón del ratón —se le señala—. Vamos a practicar un poco...»*. Se realizaron hasta tres veces los cuatro ítems de práctica, pero normalmente con una vez fue suficiente. Una vez el evaluador tuvo claro que el niño había comprendido lo que debía hacer y presentaba un correcto dominio del ratón se pasaba a la evaluación diciendo, *«bien, ahora vamos a empezar a jugar en serio. Tienes que hacer lo mismo con los ejercicios que van apareciendo en la pantalla. Recuerda que sólo hay una alternativa correcta. Lo importante es no cometer errores, pero todo el mundo comete algunos. No hay que despistarse ni entretenerse. ¿Lo has comprendido todo? Muy bien, pues empezamos»*. Durante la aplicación de la prueba el ordenador registró el número de errores y la latencia de respuesta. Una vez finalizada la prueba, se obtenían los índices de impulsividad (PI) y de eficacia (PE).

Como ya hemos indicado al exponer la prueba, los índices PI y PE de Salkind y Wright (1977) se calculan a través de las puntuaciones Z de los errores y las latencias. La PI es la resta de estas dos puntuaciones y la PE es la suma. Los cálculos se realizan por curso, de manera que los índices resultantes para cada grupo de edad presentan una media de 0, pero una desviación estándar (DE) variable, normalmente con un valor

entre 1 y 2. Muchos autores utilizan ya esa distribución para establecer los grupos y realizar las comparaciones entre sujetos, pero si son de distintos cursos, como es nuestro caso, es arriesgado. Por ejemplo, si primer curso presenta una DE de 1,01 y cuarto de 1,86 (los dos con una media de 0 porque esto si que es inalterable en todos los cursos) y se decide que se considerarán «*impulsivos*» a aquellos niños por encima de 1 DE resulta que un niño de primero con una PI de 1,10 cumple perfectamente con este criterio, pero otro niño de cuarto con la misma puntuación, aunque por definición se incluirá en el grupo de impulsivos, resultará que, dentro de su curso, realmente se halla dentro de la normalidad puesto que está a 0,76 puntos de superar el valor de 1 DE. Para subsanar este artefacto matemático hay que volver a tipificar en puntuaciones Z los índices PI y PE por curso para que en cada uno de ellos quede definitivamente una distribución de media=0 y desviación estándar=1, que permita comparar sujetos de distinta edad. Ahora sí que cuando incluyamos sujetos de +1 DE de distintos cursos se respetará ese valor dentro de su grupo de edad de referencia.

Tanto la PI como la PE, como ya hemos visto anteriormente, presentan una distribución con $M=0$ y $DE=1$. Con estas puntuaciones, por cada curso o grupo de edad, se pueden establecer diferentes sistemas para clasificar a los sujetos: los sistemas de cuadrantes y el integrado ya los hemos explicado en la primera parte de este trabajo. Ahora nos interesa destacar nuestra propuesta: el sistema que hemos denominado IMAT.

El sistema IMAT intenta subsanar algunos problemas detectados en el sistema integrado. En el integrado se propone trabajar de modo independiente o bien con la PI, o bien con la PE porque de hecho matemáticamente la correlación entre ambas variables es de cero. Sin embargo, si se decide trabajar con la PI y clasificar a los sujetos en «*impulsivos*» y «*reflexivos*» surge la duda de si es lo mismo considerar las tres posibilidades que nos podemos encontrar en un mismo grupo. Por ejemplo, dentro del grupo impulsivo, podemos encontrarnos un niño que cumpla esa condición pero que además presenten una PE tendente a la eficacia, otro con una PE tendente a la ineficacia y otro con una PE indefinida o dentro de la normalidad. ¿Cabe considerar que la impulsividad de estos tres niños

tendrá las mismas implicaciones?. Por otra parte, ¿en qué punto del continuo de puntuaciones de la PI hay que establecer la separación entre el grupo impulsivo y reflexivo? Muchos autores lo hacen a partir de la media cero, pero tampoco parece razonable que un sujeto con una PI de 2 y otro con una PI de 0,01 presenten el mismo grado y la misma consideración en su tendencia impulsiva. Para solucionar estos problemas nosotros proponemos el siguiente sistema de clasificación de sujetos:

SISTEMA IMAT	VALOR PI	VALOR PE	RESULTADO
En Estilo			Basada en la PI, ejerciendo control sobre la PE
Impulsivo	>+1	-1/+1	Tendencia a la impulsividad
Flexible/Normal	-1/+1	-1/+1	Estilo flexible, con un nivel medio de eficacia
Reflexivo	<-1	-1/+1	Tendencia a la reflexividad
En Eficacia			Basada en la PE, ejerciendo control sobre la PI
Ineficaz	-1/+1	>+1	Tendencia a la ineficacia
Normal/Flexible	-1/+1	-1/+1	Nivel medio de eficacia, con estilo flexible
Eficaz	-1/+1	<-1	Tendencia a la eficacia
Interacción			PI y PE alejados más de 1 DE de la media
Impul./Eficaz	> +1	< -1	Sujetos impulsivos y eficaces
Impul./Ineficaz	> +1	> +1	Sujetos impulsivos e ineficaces
Refle./Eficaz	< -1	< -1	Sujetos reflexivos e ineficaces
Refle./Ineficaz	< -1	> +1	Sujetos reflexivos e ineficaces

De esta manera se puede optar por trabajar con dos formas independientes de clasificación de los sujetos, basadas respectivamente en la PI y en la PE, más el complemento que supone la interacción entre ambos índices en puntuaciones «*extremas*». El punto de corte de +/- 1 DE, aunque arbitrario, tiene cierta lógica puesto que en la curva

normal incluye aproximadamente al 68% de los sujetos de la población. Estos sujetos los consideramos de estilo flexible/eficacia normal y los diferenciamos de los claramente tendentes hacia un estilo o hacia un nivel de eficacia. Cuando la versión de la prueba de evaluación presente una elevada correlación negativa entre errores y latencias, como es el caso de la TEF según veremos en el análisis de resultados, cabe esperar que la inmensa mayoría de sujetos estén en el grupo flexible/normal, impulsivo o reflexivo y sólo una minoría estén en los grupos eficaz e ineficaz. Por su parte, los grupos de interacción se constituirán, según nuestra hipótesis, por unos porcentajes aún menores de sujetos, puesto que se trata de verdaderas excepciones a la regla.

En definitiva, el objetivo es poder establecer comparaciones en variables de rendimiento y comportamiento en el aula entre impulsivos y reflexivos sin que su PI se halle «contaminada» con una puntuación extrema en la PE que pueda alterar los resultados. Con el sistema IMAT sólo cuando el nivel de eficacia sobre la TEF se halla en la normalidad se comparan estos dos estilos de procesar la información. Y este es un hecho, a nuestro entender, bastante novedoso en la investigación en el campo de la reflexividad-impulsividad.

En el caso de **la Tarea de vigilancia (TVI)** el procedimiento fue el siguiente. Al niño se le explicaba que *«en la pantalla del ordenador irán apareciendo números bastante rápidos. Lo que tienes que hacer es apretar la barra de espacio cada vez que veas un 6 y justo después un 3»*. Durante dos minutos el niño practica este ejercicio y el evaluador se cerciora de que realmente ha comprendido lo que tiene que hacer. Posteriormente se le informa de que *«bien, ahora vamos a empezar en serio. Procura no despistarte y trabaja hasta que el ordenador te lo indique. Recuerda, sólo tienes que apretar la barra de espacio cuando veas un 6 y justo después un 3»*. Como ya se ha explicado, al final de cada aplicación el ordenador cinco puntuaciones directas, de las cuáles para análisis posteriores nosotros hemos escogido las tres más importantes: los aciertos, los errores de comisión y el tiempo de reacción a los aciertos.

Tal y como hemos desarrollado con cierta profusión en la primera parte del trabajo, la teoría de detección de señales nos permite a su vez la obtención de dos índices combinados, a partir de las puntuaciones de aciertos y comisiones: el índice de sensibilidad y el sesgo de respuesta. En el primer caso hemos optado, como la mayoría de autores, por la d' , esto es, la resta entre la puntuaciones Z de aciertos y comisiones. En cambio, en el segundo caso, al contrario de muchos autores que optan por el índice Beta, nosotros hemos preferido el criterio C que, como hemos explicado en su momento, visualmente y por procedimiento de cálculo parece más sencillo, si bien los dos ofrecen exactamente la misma información, es decir, el posible sesgo de respuesta hacia el conservadurismo (predominio de errores de omisión) o el riesgo (predominio de errores de comisión) del sujeto.

Tanto d' como C funcionan con puntuaciones tipificadas Z y, por tanto, su media es 0 y su DE es 1, calculándose dentro de cada curso en particular. Si bien hay que proceder del mismo modo que anteriormente se ha explicado en el caso de los índices PI y PE de la TEF para garantizar que se puedan hacer comparaciones entre sujetos de distintas edades. El índice de sensibilidad es el más importante, como además pronto se verá refrendado en el análisis de resultados, y expresa en cierto modo la capacidad de atención sostenida del sujeto. Sin duda para poder interpretar correctamente esta capacidad es necesario el control del sesgo de respuesta: pero ya por sí solo d' da una información mucho más rica de lo que es trabajar por separado solamente con aciertos y errores de comisión. Hay que tener en cuenta que si tomamos sólo en consideración los aciertos de un sujeto, y éstos son elevados, no podemos decir gran cosa sobre su capacidad atencional, porque pudo suceder que cometiera muchísimos errores de comisión que le permitiera, a su vez, el conseguir muchos aciertos por azar. Al revés, el que un sujeto presente pocos errores de comisión nos da una idea general de cierta capacidad de control motor, pero no de su capacidad atencional que dependerá de los aciertos. El índice de sensibilidad pone en relación ambas puntuaciones y soluciona satisfactoriamente la cuestión.

Para los análisis sobre las variables del ECRA propondremos dos sistemas de clasificación de los sujetos, a partir del índice de sensibilidad. Un primer sistema, más general, simplemente define como sujetos de «*elevada capacidad atencional*» a aquéllos que presentan una d' mayor de +1, mientras definimos a los de «*baja capacidad atencional*» a aquéllos con una puntuación menor de -1, siendo el intervalo alrededor del 0 definitorio de sujetos con capacidad *normal*. Un segundo sistema establece una distinción más amplia por grupos. En este caso entre la puntuación +1 y +1,5 está el grupo de *alta capacidad*, y por encima de +1,5 el de *muy alta capacidad*. Por su parte, entre -1 y -1,5 está el grupo de *baja capacidad*, y por debajo de -1,5 el de *muy baja capacidad*. En el análisis de resultados se irán presentando los datos obtenidos por cada sistema.

En **el caso de la ECRA** sus subescalas se han utilizadas como variables criterio para analizar la influencia de las variables predictivas: la reflexividad-impulsividad y la atención sostenida, sobre variables de comportamiento y rendimiento en el aula. La escala fue cumplimentada por los maestros tutores de los niños.

El análisis de resultados se basa, en primer lugar, en los resultados descriptivos y comparativos por cursos y sexos de la TEF y la TVI. A continuación se analizan los datos de fiabilidad disponibles de las dos pruebas, teniendo presente que en el caso de la TVI únicamente es posible trabajar con procedimientos de test-retest por la propia naturaleza de la prueba. En el apartado siguiente se aborda la relación entre las puntuaciones directas y los índices del estilo impulsivo y la atención sostenida. Y, finalmente, aunque por separado, se abordan las implicaciones sobre las subescalas de la ECRA de ambas dimensiones. Para el análisis de datos se ha utilizado el paquete SPSS y únicamente hay que aclarar que en los contrastes a posteriori de los análisis de la variancia se ha utilizado la prueba de Scheffé.

6. ANÁLISIS DE LOS RESULTADOS.

6.1 Resultados descriptivos y comparación por sexo y por cursos.

6.1.1 Datos de la TEF-IMAT.

En la tabla 6 se comparan los resultados por sexos en las medidas de la Tarea de Emparejamiento de Figuras (TEF). El número de sujetos se puede consultar en la tabla 5.

Tabla 6. Comparación de los resultados de la TEF por sexos

Curso	Niños		Niñas		t	
	Media	DE	Media	DE		
1	Errores	32,28	8,25	29,81	8,12	2,24*
	Latencia	7,70	3,24	9,18	3,67	3,26**
	PI	0,33	1,63	-0,45	1,81	3,28**
	PE	0,01	1,01	0,14	0,87	ns
2	Errores	24,55	8,60	23,09	8,86	ns
	Latencia	10,12	3,79	10,54	3,77	ns
	PI	0,15	1,75	-0,15	1,87	ns
	PE	0,01	0,84	0,00	0,95	ns
3	Errores	18,66	8,04	19,04	7,23	ns
	Latencia	11,31	4,79	10,65	3,99	ns
	PI	0,00	1,76	0,00	1,61	ns
	PE	0,10	0,90	-1,14	0,83	ns
4	Errores	14,61	8,18	14,09	6,79	ns
	Latencia	14,01	5,45	13,50	5,37	ns
	PI	-0,18	1,99	0,00	1,69	ns
	PE	0,00	0,95	0,01	0,79	ns

* $p < .05$, ** $p < .01$

PI (índice de estilo impulsivo), PE (índice de eficacia)

Los resultados indican que prácticamente no hay diferencias por sexos en ninguna variable. La excepción la constituye el primer curso donde las niñas se muestran significativamente más reflexivas que los niños (menos errores y latencias más largas), aunque en la práctica la diferencia es mínima. A la vista de estos resultados los siguientes análisis se realizan para la muestra total, sin diferenciar entre sexos.

En la tabla 7 se pueden consultar las comparaciones por cursos sobre las medidas de la TEF.

Tabla 7. Comparación de los resultados de la TEF por cursos

	1°	2°	3°	4°	F	Contrastes
Errores						
Media	31,38	23,90	18,82	14,37	159,75**	1-2-3-4
DE	8,27	8,72	7,69	7,55		
Latencias						
Media	8,22	10,31	11,04	13,77	57,84**	1-2,3-4
DE	3,46	3,78	4,48	5,40		

* $p < .05$, ** $p < .01$

Los resultados indican diferencias muy importantes en función del curso, es decir, en función de la variable edad. En el caso de los errores las diferencias son estadísticamente significativas entre los cuatro cursos, con tendencia a una clara disminución. En el caso de las latencias estadísticamente se producen tres grupos: primer curso, segundo y tercero (entre ellas no se llega a la significación) y cuarto. La tendencia es a aumentar la latencia de respuesta. Dado que el índice de estilo, PI, y el de eficacia, PE, se calcula dentro de cada curso no tiene sentido hacer comparaciones entre ellos, pero sí cabe señalar que con la edad se produce una clara tendencia al aumento de la reflexividad.

6.1.2 Datos de la TVI-IMAT.

En la tabla 8 se comparan los resultados por sexo en la tarea de vigilancia (TVI), prueba de evaluación de la capacidad de atención sostenida.

Tabla 8. Comparación de los resultados de la TVI por sexos

Curso	Niños		Niñas		t		
	Media	DE	Media	DE			
1	d'	0,01	1,57	-0,01	1,57	2,35*	
	C	-0,09	0,58	0,13	0,67		
	Aciertos	58,63	11,76	56,01	11,65		
	Comisiones	44,53	22,63	39,66	25,53		
	T. Reacción	384,71	65,93	431,9	52,47		5,27**
2	d'	0,15	1,56	-0,2	1,57	3,06**	
	C	-0,15	0,59	0,21	0,6		
	Aciertos	70,32	8,44	65,32	9,58		3,34**
	Comisiones	27,27	14,63	24,63	13,24		
	T. Reacción	371,74	71,28	395,91	64,19		2,10*
3	d'	-0,22	1,46	0,3	1,78	3,72**	
	C	0,05	0,57	-0,07	0,61		
	Aciertos	70,15	8,53	73,67	9,76		
	Comisiones	28,68	15,83	26,3	18,77		
	T. Reacción	331,08	51,99	370,66	54,07		
4	d'	-0,08	1,76	0,09	1,73	4,13**<	
	C	-0,07	0,47	0,08	0,51		
	Aciertos	77,2	8,15	76,69	8,82		
	Comisiones	19,42	11,84	16,78	11,11		
	T. Reacción	342,44	55,11	382,16	59,98		

* $p < .05$, ** $p < .01$

d' (índice de sensibilidad), C (sesgo de respuesta)

En general no se aprecian diferencias por sexos, con algunas excepciones. En primer y segundo curso los niños muestran un sesgo de respuesta significativamente más *arriesgado* que las niñas (con más errores de comisión, falta de control motor). En cualquier caso todos los valores están alrededor de la media cero. En segundo curso se da el único caso en que una de las puntuaciones directas, los aciertos, diferencian significativamente a ambos sexos, en el sentido de una mejor puntuación de los niños. Por último, y éste sí que es el resultado más constante, en todos los cursos, en mayor o menor grado, el tiempo de respuesta ante los aciertos de los niños es significativamente más corto que el de las niñas. En todo caso como esta medida no se utiliza en los índices combinados de sensibilidad y sesgo de respuesta no se realizarán más análisis comparativos por sexo, dado que en realidad las diferencias son mínimas.

En la tabla 9 se presentan los datos comparativos por cursos de las principales puntuaciones directas de la tarea de vigilancia.

Los resultados indican claras diferencias por cursos en todas las variables. Las más evidentes se refieren a la puntuación de aciertos, en la que los cuatro cursos se diferencian estadísticamente y la tendencia in-

Tabla 9. Comparación de los resultados de la TEF por cursos

	1°	2°	3°	4°	F	Contrastes
Aciertos						
Media	57,59	68,20	71,64	76,96	112,74**	1-2-3-4
DE	11,76	9,25	9,19	8,44		
Comisiones						
Media	42,59	26,15	27,68	18,18	55,19**	1-2,3-4
DE	23,88	14,07	17,08	11,54		
T. Reacción						
Media	403,53	381,91	347,93	361,05	21,46**	1-2,3-3,4-4
DE	65,04	69,20	56,17	60,60		

* $p < .05$, ** $p < .01$

dica a un aumento de su número en función de la edad. En la puntuación de errores de comisión las diferencias más evidentes afectan al primer curso en relación al cuarto, mientras segundo y tercero constituyen un grupo diferenciado de estos dos. En cualquier caso, la tendencia, es con el aumento de edad se reduzca significativamente el número de los errores. Finalmente en el tiempo de reacción las diferencias son de menor magnitud, si bien también indican una disminución significativa de esta variable bajo la influencia de la edad.

A partir del sistema de cálculo del índice de sensibilidad (d') y del sesgo de respuesta (criterio C) propuesto en el programa de Reales y Ballesteros (1997), la tabla 10 muestra la tendencia de ambas medidas por curso a partir de los porcentajes de aciertos y de errores de comisión.

Tabla 10. Porcentajes de aciertos y errores de comisión y evolución del índice de sensibilidad (d') y el sesgo de respuesta (C) por cursos

	Curso			
	1	2	3	4
% Aciertos	0,64	0,76	0,80	0,86
% errores de comisión	0,20	0,12	0,13	0,09
d'	1,20	1,88	1,97	2,42
C	0,24	0,23	0,14	0,13

En primer lugar cabe destacar que todos los cursos presentan una media de sensibilidad superior al punto sobre la media cero. En segundo lugar la evolución es en la dirección esperada, es decir, con la edad se produce un aumento muy significativo de la sensibilidad (la capacidad de atención sostenida) del niño. Por lo que respecta al sesgo de respuesta, en primer lugar cabe señalar que todas las medias están alrededor de la media cero, es decir, sin sesgo de respuesta definido aunque tendente a ser conservador (puntuaciones C positivas). En segundo lugar la evolución por edades indica que el poco sesgo que se detecta en primero va todavía reduciéndose más en cursos superiores,

con lo que el niño va ganando control sobre su respuesta motora. Hay que recordar que el hecho de encontrar tan poco sesgo entre los sujetos puede deberse al hecho de que en la muestra se han eliminado sujetos extremos para las variables aciertos y errores de comisión.

6.2 Análisis de la fiabilidad y consistencia de las medidas

6.2.1 Datos de la TEF-IMAT

En la tabla 11 se presentan los datos de fiabilidad obtenidos para la Tarea de Emparejamiento de Figuras. Los datos para el sistema test-retest provienen de una muestra mucho más reducida de sujetos ($n=68$) que los de los otros procedimientos ($n=771$).

Tabla 11. Datos de fiabilidad por consistencia interna y test-retest de la TEF

	1er curso		2º curso		3er curso		4º curso	
	Errores	Latencias	Errores	Latencias	Errores	Latencias	Errores	Latencias
Alfa	,83	,91	,81	,90	,77	,88	,77	,89
Guttman	,80	,83	,79	,83	,73	,81	,73	,81
Test-retest	,69	,55	,65	,68	,64	,97	,64	,74

Nota: los datos de los coeficientes Alfa de Cronbach y Dos Mitades de Guttman se obtuvieron sobre la muestra del estudio ($n=771$), mientras los del Test-Retest provienen de una muestra de 68 sujetos, por cursos respectivamente 20, 19, 11 y 18.

Los resultados señalan una muy aceptable fiabilidad y consistencia interna tanto para los errores como para las latencias. En el caso de los errores el coeficiente alfa de Cronbach fluctúa entre .77 y .83, mientras en las latencias no baja del .88. Por su parte, en el caso del coeficiente de Guttman los valores son similares, aunque algo inferiores en su conjunto, de nuevo favorables a las latencias. Por lo que respecta al procedimiento test-retest los valores, como suele ser habitual, son los más bajos en su conjunto, y más favorables a los errores que a las latencias. Los errores fluctúan entre .64 y .69 y las latencias entre .55 y .97.

6.2.2 Datos de la TVI-IMAT

En la tabla 12 se pueden observar los datos de fiabilidad obtenidos para la Tarea de Vigilancia en una muestra de 73 sujetos.

Tabla 12. Datos de fiabilidad test-retest para la TVI

	1er curso			2º curso			3er curso			4º curso		
	acierto	comis	tr	acierto	comis	tr	acierto	comis	tr	acierto	comis	tr
test												
-retest	,77	,56	,73	,69	,60	,66	,73	,62	,85	,79	,88	,81

n=73 (respectivamente por curso 19, 18, 18 y 18). Acierto (total aciertos), comis (errores de comisión), tr (tiempo de reacción a los aciertos).

Los resultados fluctúan, en el caso de la puntuación de aciertos, entre .69 y .79. En el caso de los errores de comisión son algo más bajos y más variables, entre .56 y .88. Por último en la medida del tiempo de reacción están entre .66 y .85. En general los valores de fiabilidad por consistencia interna parecen razonablemente elevados.

6.3 La relación entre el estilo impulsivo y la falta de atención sostenida

En la tabla 13 aparece la matriz de correlaciones entre las puntuaciones directas de la Tarea de Emparejamiento de Figuras y la Tarea de Vigilancia.

Como se indica en la propia tabla se ha optado por destacar las correlaciones iguales o superiores al .30 puesto que, aunque en valores muy inferiores, dado el tamaño de la muestra, ya se consiguen niveles de significación estadística, se ha creído oportuno partir de un punto de corte relativamente relevante en la práctica educativa. En primer lugar, centrándonos en las correlaciones intra dimensión, cabe destacar la correlación entre errores y latencias de la TEF puesto que es uno de los primeros índices de validez de constructo de la reflexividad-impulsividad. Como vemos los valores son relativamente elevados y negativos, como era de desear: fluctúan entre -.51 y -.65, con un valor de -.67 para la muestra to-

Tabla 13. Coeficientes de correlación entre las puntuaciones directas de la TEF y de la TVI por cursos

	Errores	Latencias	Aciertos	Comisiones	T. Reacción
1° (n=177)					
Errores	1,00				
Latencia	-,51	1,00			
Aciertos	-,19	,00	1,00		
Comisiones	,16	-,08	-,22	1,00	
T. Reacción	-,06	,07	-,30	-,17	1,00
2° (n=135)					
Errores	1,00				
Latencia	-,64	1,00			
Aciertos	-,29	,14	1,00		
Comisiones	,08	-,11	-,23	1,00	
T. Reacción	-,10	,30	-,24	-,26	1,00
3° (n=93)					
Errores	1,00				
Latencia	-,56	1,00			
Aciertos	-,22	,22	1,00		
Comisiones	,44	-,38	-,31	1,00	
T. Reacción	-,16	,20	-,02	-,23	1,00
4° (n=139)					
Errores	1,00				
Latencia	-,65	1,00			
Aciertos	-,40	,20	1,00		
Comisiones	,35	-,25	-,51	1,00	
T. Reacción	-,10	,17	,04	-,19	1,00
Total (n=584)					
Errores	1,00				
Latencia	-,67	1,00			
Aciertos	-,56	,36	1,00		
Comisiones	,45	-,35	-,47	1,00	
T. Reacción	,13	,03	-,30	-,04	1,00

Nota 1: se resaltan aquellas correlaciones iguales o superiores al .30 (dado el tamaño de la muestra es posible encontrar correlaciones significativas en valores cercanos a .10 pero en la realidad se consideran muy bajas).

Nota 2: El número de sujetos que aparece se refiere a la muestra total sin sujetos extremos para ninguna de las medidas de la TVI y la TEF. De todos modos, para el caso concreto de la correlación errores/latencias se han podido mantener más sujetos y la muestra es la siguiente: 1° (n=243), 2° (n=193), 3° (n=120), 4° (n=168), total (n=724).

tal. En cuanto a las medidas de la TVI, los aciertos ofrecen correlaciones negativas importantes básicamente con los errores de comisión, aunque sólo en los dos últimos cursos (en cuarto el índice llega a $-.51$). La relación entre los aciertos y el tiempo de reacción, aunque en general es en la dirección esperada (tiempos más cortos para los niños con mayores aciertos) es altamente variable, y aunque en la muestra total llega al $-.30$, las diferencias entre cursos son muy grandes. Por tanto, pues, las tendencias son, en la TEF, a mayor latencia de respuesta menos errores, mientras en la TVI, especialmente en los últimos cursos, el mayor número de aciertos implica (aún siendo medidas independientes) menos errores de comisión y, en menor grado, una reducción del tiempo de reacción.

En segundo lugar, y en lo que respecta a las relaciones entre dimensiones, observamos que los errores de la TEF correlacionan más que las latencias sobre las variables de la TVI. La tendencia es a obtener correlaciones más elevadas en los cursos superiores, aún con alguna excepción. En la muestra total se dibujan claramente las tendencias: los errores llegan a niveles de correlación muy significativos con los aciertos ($-.56$) y los errores de comisión ($.45$), mientras las latencias también lo hacen, aunque sin tanta intensidad ($.36$ con los aciertos y $-.35$ con los errores de comisión). En general, pues, cabe sospechar la presencia de una relación entre las dimensiones reflexividad-impulsividad y atención sostenida. En la tabla 14 se analiza esta posibilidad esta vez no a partir de las puntuaciones directas, sino de los índices combinados para ambas variables.

Los resultados indican ciertamente la presencia de una correlación moderada entre ambas dimensiones, si bien claramente influenciada por la edad. La tendencia general afecta casi en exclusiva a la PI (puntuación de estilo) y a d' (índice de sensibilidad) en la siguiente dirección: los niños más reflexivos (especialmente de tercer y cuarto curso) tienden a presentar una mejor atención sostenida (o viceversa). Por su parte, las correlaciones que afectan a PE y a sesgo de respuesta son nulas o poco relevantes.

En la tabla 15 se realiza un nuevo análisis de la relación entre la reflexividad-impulsividad y la atención sostenida a través de procedimientos de comparación de medias. En este caso se han obtenido grupos de sujetos impulsivos/reflexivos y eficaces/ineficaces, tanto por el procedi-

Tabla 14. Coeficientes de correlación entre los índices de la reflexividad-impulsividad y de la atención sostenida

Curso	N		PI	PE		PI	PE
1	177	<i>d'</i>	-,16	-,18	C	-,02	,08
2	135	<i>d'</i>	-,22	-,07	C	,11	,15
3	93	<i>d'</i>	-,44	-,03	C	-,17	-,05
4	139	<i>d'</i>	-,38	-,17	C	,00	,11
total	544	<i>d'</i>	-,28	-,12	C	-,01	,08

PI: índice de estilo reflexividad-impulsividad (a mayor puntuación, mayor impulsividad)

PE: índice de eficacia-ineficacia (a mayor puntuación, mayor ineficacia)

d': índice de sensibilidad (a mayor puntuación, mayor atención sostenida)

C: sesgo de respuesta conservador-arriesgado (a mayor puntuación sesgo más conservador)

Tabla 15. Comparación entre los grupos clasificados a partir de la dimensión reflexividad-impulsividad sobre los índices de atención sostenida (sensibilidad y sesgo de respuesta)

		Índice de Sensibilidad (<i>d</i>)		Sesgo de Respuesta (C)	
		M	DE	M	DE
Sistema integrado	<i>n</i>				
<i>Impulsivos</i>	248	-0,42	1,57	0,02	0,62
<i>Reflexivos</i>	296	0,28	1,60	0,01	0,57
<i>t</i>		-5,17**		0,62	
<i>Ineficaces</i>	261	-0,21	1,63	0,03	0,64
<i>Eficaces</i>	283	0,12	1,59	-0,03	0,54
<i>t</i>		-2,35*		1,21	
Sistema IMAT (estilos)					
<i>Impulsivos</i>	109	-0,59	1,57	0,07	0,67
<i>Reflexivos</i>	145	0,53	1,69	0,05	0,54
<i>t</i>		-5,37**		-1,63	
Sistema IMAT (eficacia)					
<i>Ineficaces</i>	23	-0,36	1,57	0,10	0,59
<i>Eficaces</i>	48	0,19	1,41	-0,11	0,54
<i>t</i>		-1,46		1,55	

(continúa)

(continuación)

Tabla 15. Comparación entre los grupos clasificados a partir de la dimensión reflexividad-impulsividad sobre los índices de atención sostenida (sensibilidad y sesgo de respuesta)

		Índice de Sensibilidad (d')		Sesgo de Respuesta (C)	
Sistema IMAT (interac.)					
<i>Impulsiv Eficac</i> (1)	7	-0,50	1,95	-0,40	0,87
<i>Reflexiv Eficac</i> (2)	16	0,73	1,18	-0,01	0,54
<i>Reflexiv Inefica</i> (3)	32	0,05	1,54	-0,10	0,68
<i>Impulsiv Inefica</i> (4)	18	-0,93	1,21	0,22	0,68
F		4,08* (2,3,1-1,3,4)		1,66	

* $p < .05$, ** $p < .01$

Sistema integrado: impulsivos $PI > 0$ y reflexivos $PI < 0$. Ineficaces $PE > 0$ y eficaces $PE < 0$.

Sistema IMAT: Impulsivos $PI > +1$ y reflexivos $PI < -1$ (los dos grupos con una PE entre ± 1 , es decir, en la *normalidad*). Ineficaces $PE > +1$ y eficaces $PE < -1$ (los dos grupos con una PI entre ± 1 , es decir, en la *normalidad*). Grupos de interacción: sujetos con PI y PE de ± 1 desviación estándar, es decir, las dos puntuaciones fuera de la *normalidad*:

miento de clasificación integrado de Salkind y Wright (1977), como por el sistema IMAT, y se han comparado sus resultados en sensibilidad (d') y sesgo de respuesta (C).

En primer lugar, y como era lógico dado sus bajos niveles de correlación, no aparece ninguna comparación significativa relacionada con el sesgo de respuesta y, por tanto, se puede descartar la relación entre este índice y la R-I. En lo que respecta a la sensibilidad o capacidad de atención sostenida se confirma que las correlaciones moderadamente elevadas antes detectadas llegan a producir muchas diferencias entre los distintos subgrupos clasificados. Por ejemplo, en el sistema integrado (que simplemente divide a los grupos por encima y por debajo de la media cero de sus índices) se observa una clara superioridad de los sujetos reflexivos sobre los impulsivos en atención sostenida. En el otro índice, la PE , también se observa la superioridad de los sujetos eficaces sobre los ineficaces aunque la magnitud de la diferencia es bastante inferior.

En el sistema de clasificación IMAT los grupos se perfilan mucho más, tal y como explica el pie de tabla. En este caso, y aún viéndose los

grupos mucho más reducidos dado que se seleccionan a los sujetos claramente impulsivos y reflexivos sin que se diferencien en la PE, se mantienen prácticamente igual las diferencias amplias a favor de los reflexivos en atención sostenida. En cambio en la comparación a través del otro índice las diferencias han desaparecido; ahora los sujetos claramente eficaces e ineficaces (controlando su PI) no se diferencian estadísticamente entre sí en función de la capacidad de atención sostenida. En los subgrupos de interacción (muy reducidos porque se trata en todos los casos de puntuaciones extremas combinadas entre PI y PE) cabe destacar que la principal diferencia se da entre el subgrupo de reflexivos-eficaces y el de impulsivos-ineficaces, con diferencias claramente favorables a los primeros. En general de nuevo parece ser el índice de estilo y no el de eficacia el que marca la pauta diferencial en atención sostenida.

6.4 Las implicaciones de la impulsividad cognitiva en el aula

En la tabla 16 aparecen las correlaciones entre las medidas directas de la TEF y los índices de la dimensión reflexividad-impulsividad (PI y PE) sobre las variables de rendimiento y comportamiento evaluadas a través de la Escala de Conducta y Rendimiento en el Aula para maestros (ECRA).

La tendencia general más evidente indica que las medidas de la R-I se reflejan especialmente sobre tres subescalas: la de rendimiento académico, la de problemas de trabajo escolar y la de inatención, mientras las correlaciones con las de impulsividad y problemas de interacción social son mucho menores. A pesar de que hay variaciones más o menos importantes por cursos, en la muestra total se observa la relación moderada pero muy significativa entre los errores y la PI sobre el rendimiento académico: a mayor estilo impulsivo pero rendimiento académico, si bien la correlación dependen mucho más de los errores (-.30) que de las latencias (.18). La variable «*problemas de trabajo escolar*» presenta prácticamente la misma tendencia, con niveles de correlación casi equivalentes. Por su parte la variable «*inatención*», aún siguiendo la misma tendencia, presenta correlaciones inferiores al .30. Llama la atención, valga la redundancia, que en este caso han «*fallado*» las correlaciones de

Tabla 16. Análisis de la relación entre las medidas de la TEF y las variables de rendimiento y comportamiento en el aula (ECRA)

Curso		R.Acad.	Prob.Escolar	Inatención	Impulsividad	Prob.Social
1 (n=243)						
	PI	-,31	,30	,21	,09	,10
	PE	-,15	,04	,09	-,06	,06
	Errores	-,34	,28	,23	,05	,11
	Latencia	,19	-,24	-,14	-,11	-,05
2 (n=193)						
	PI	-,37	,33	,29	,22	,24
	PE	-,22	,31	,23	,09	,11
	Errores	-,43	,43	,36	,24	,27
	Latencia	,24	-,17	-,17	-,16	-,17
3 (n=120)						
	PI	-,43	,37	,33	,2	,20
	PE	-,27	,38	,35	,32	,38
	Errores	-,49	,49	,44	,32	,34
	Latencia	,26	-,15	-,14	-,03	,00
4 (n=168)						
	PI	-,26	,22	,11	,14	,17
	PE	-,21	,15	,12	,12	,19
	Errores	-,31	,25	,15	,17	,23
	Latencia	,15	-,14	-,05	-,08	-,07
Total (n=724)						
	PI	-,34	,30	,24	,15	,16
	PE	-,20	,20	,18	,08	,15
	Errores	-,30	,30	,21	,16	,21
	Latencia	,18	-,17	-,10	-,10	-,10

**Nota: se resaltan aquellas correlaciones iguales o superiores al .30 (dado el tamaño de la muestra es posible encontrar correlaciones significativas en valores cercanos a .10 pero en la realidad se consideran muy bajas).*

cuarto curso que, al contrario de lo que sucede en prácticamente todo el estudio, registran aquí una bajada importante con respecto a los cursos precedentes. En resumen, el estilo impulsivo se relaciona fundamentalmente con medidas ligadas al rendimiento académico y al aprendizaje y, aunque la aportación de las latencias no es despreciable, las correlaciones dependen en mayor grado de los errores sobre la TEF.

En la tabla 17 aparecen las comparaciones entre los distintos subgrupos derivados de la R-I sobre las medidas de la ECRA. En este caso se contempla el sistema de cuadrantes, esto es, la doble división por las medianas de errores y latencias, y el sistema integrado, esto es, la clasificación a partir de los índices combinados PI y PE.

En los dos sistemas tradicionales de clasificación de sujetos a partir de la R-I aparecen diferencias significativas en todas las subescalas de la ECRA, aunque de distinta magnitud. Como era de esperar a partir de

Tabla 17. Comparación entre impulsivos y reflexivos en los sistemas de clasificación tradicionales en las variables de rendimiento y comportamiento en el aula (ECRA)

Curso	n	R.Acad.		Prob.Escolar		Inatención		Impulsividad		Prob.Social		
		M	DE	M	DE	M	DE	M	DE	M	DE	
Sistema cuadrantes												
Impulsivos	249	9,12	2,94	10,12	4,10	18,73	7,51	12,72	5,68	8,57	4,36	
<i>Reflexivos</i>	212	10,53	2,57	8,34	3,49	15,06	6,96	10,86	4,71	7,23	3,30	
t		-5,42**		4,98**		5,42**		3,78**		3,69**		
Sistema integrado												
Impulsivos	342	9,15	2,94	10,16	4,09	18,54	7,48	12,49	5,49	8,44	4,30	
<i>Reflexivos</i>	310	10,40	2,59	8,31	3,49	15,25	6,91	10,92	4,73	7,42	3,44	
t		-5,68**		6,18**		5,81**		3,88**		3,30**		

* $p < .05$, ** $p < .01$

Sistema de cuadrantes: impulsivos (por encima de la mediana de errores y por debajo de la de latencias) y reflexivos (por debajo de la mediana de errores y por encima de la de latencias), con sujetos rápidos-exactos y lentos-inexactos eliminados.

Sistema integrado: impulsivos, por encima de la puntuación 0 en la PI, y reflexivos, por debajo de la puntuación 0 en la PI (no se elimina a ningún sujeto).

las correlaciones anteriormente observadas, las diferencias son mucho más ostensibles en rendimiento, en problemas de trabajo escolar y en inatención, y mucho menos en problemas de impulsividad y de interacción social. Las diferencias siempre indican mejores puntuaciones de rendimiento y trabajo académico, así como menores problemas de comportamiento, para los niños reflexivos. La única diferencia entre los dos sistemas de clasificación es que el integrado consigue mayores valores de t en casi todas las comparaciones.

En la tabla 18 aparecen las mismas correlaciones pero ahora aplicadas al sistema IMAT de clasificación de sujetos. En este caso cuando se obtienen los grupos en estilo se exigen puntuaciones de corte de ± 1 desviación estándar, mientras se controla la PE. Lo mismo sucede, pero a la inversa, cuando se obtienen los grupos en eficacia. Y, finalmente, se contempla la posibilidad de la interacción, es decir, PI y PE por encima y por debajo de 1 desviación estándar. Lo más interesante del sistema IMAT es que tanto en la PI como en la PE contempla la presencia del grupo «flexible», esto es, el contingente de sujetos con puntuaciones PI y PE entre 0 y ± 1 desviación estándar, a los cuáles consideramos al mismo tiempo de estilo indefinido y de eficacia normal. El hecho que la TEF registre correlaciones entre errores y latencias negativas y bastante elevadas ha provocado que la mayoría de sujetos puedan diferenciarse en estilo y sólo unos pocos en eficacia o en interacción.

En el caso de la clasificación en estilo se observa la superioridad relativa de los reflexivos, pero no sólo sobre los impulsivos, sino también sobre los flexibles. Es decir, en todas las variables impulsivos y flexibles constituyen un subgrupo con puntuaciones inferiores a los reflexivos; tanto en variables de rendimiento y trabajo escolar como de comportamiento. De todos modos, de nuevo la magnitud de las diferencias es mucho mayor en las variables académicas y en la subescala de inatención (en las otras dos la significación sólo es con una $p < .05$). En el caso de la clasificación en eficacia la tendencia es la misma, aunque hay más variabilidad entre los subgrupos resultantes y desaparecen las diferencias en la medida de impulsividad; en general, de todos modos, los niños eficaces rinden mejor y tienen menos problemas de comportamiento que

Tabla 18. Comparación de los grupos clasificados según el sistema IMAT en las variables de rendimiento y comportamiento en el aula (ECRA)

Curso	R.Acad.		Prob.Escolar		Inatención		Impulsividad		Prob.Social		
	<i>n</i>	M	DE	M	DE	M	DE	M	DE	M	DE
<i>Clas. en Estilo (PI)</i>											
Impulsivos(1)	170	9,25	2,77	9,98	4,07	18,50	7,33	12,52	5,63	8,34	4,25
<i>Flexibles(2)</i>	174	9,55	2,83	9,53	3,82	17,14	7,39	11,79	5,11	8,18	3,97
<i>Reflexivos(3)</i>	138	10,70	2,46	8,31	3,46	14,70	6,84	10,76	4,73	7,14	3,19
<i>F</i>		11,50**	(1,2-3)	7,61**	(1,2-3)	10,67**	(1,2-3)	4,30*	(1,2-3)	4,18*	(1,2-3)
<i>Clas. en Efic. (PE)</i>											
<i>Ineficaces (1)</i>	28	8,04	2,62	10,79	3,46	20,07	6,39	11,71	4,44	8,57	4,01
<i>Flexibles (2)</i>	174	9,55	2,83	9,53	3,82	17,14	7,39	11,79	5,11	8,18	3,97
<i>Eficaces (3)</i>	58	11,11	2,80	7,38	3,75	13,66	6,94	10,16	4,69	6,38	3,15
<i>F</i>		12,11**	(1-2-3)	9,94**	(1,2-3)	8,60**	(1,2-2,3)	2,44		5,47**	(1,2-3)
<i>Interacción PI/PE</i>											
<i>Impulsiv</i>	8	11,38	3,38	7,13	3,48	15,38	6,99	10,88	3,87	6,75	3,15
<i>Eficac (1)</i>											
<i>Reflexiv</i>	15	10,13	2,59	8,20	4,13	15,60	7,09	12,60	6,10	7,73	3,59
<i>Eficac (2)</i>											
<i>Reflexiv</i>	31	10,39	2,38	8,16	3,35	16,03	6,73	11,06	4,23	7,32	3,40
<i>Inefica (3)</i>											
<i>Impulsiv</i>	30	7,17	2,65	12,90	3,24	23,47	6,06	15,07	5,76	11,73	4,63
<i>Inefica (4)</i>											
<i>F</i>		10,63**	(1,2,3-4)	12,92**	(1,2,3-4)	8,71**	(1,2,3-4)	3,48*		8,21**	(1,2,3-4)

* $p < .05$, ** $p < .01$

Sujetos flexibles: sujetos en la «normalidad», es decir, con una PI y una PE entre +/- 1 desviación estándar (DE).

Impulsivos y reflexivos: respectivamente una PI + 1 DE y una PI - 1DE, con la PE en la «normalidad»

Ineficaces y eficaces: respectivamente una PE + 1 DE y una PE - 1DE, con la PI en la «normalidad»

Grupos de interacción: sujetos con PI ó PE de +/- 1 DE, es decir, las dos fuera de la «normalidad».

los ineficaces. Finalmente, entre los cuatro subgrupos menores resultantes de la interacción P/PE las diferencias están entre, por un lado, «*impulsivos ineficaces*» y, por otro lado, los otros tres subgrupos. Sin duda, la interacción provoca efectos sobre la clasificación R-I puesto que es fácil darse cuenta que los sujetos «*impulsivos ineficaces*», como ya era de suponer, presentan las peores puntuaciones (en cualquiera sea el sistema de clasificación con el que queramos establecer comparaciones). Por su parte, el otro grupo extremo, los «*impulsivos eficaces*» obtienen las mejores puntuaciones. No olvidemos, de todos modos, que de 724 sujetos, una vez eliminados *outliers*, sólo 8 han entrado en este grupo.

En cualquier caso, y a modo de resumen, parece claro que con independencia del sistema de clasificación utilizado el estilo cognitivo impulsivo se asocia principalmente con mayores problemas académicos y con conductas de falta de atención (mientras es más difícil llegar a la significación en las conductas impulsivas y los problemas de interacción social).

6.5 Las implicaciones de la atención sostenida en el aula

En la tabla 19 aparecen las correlaciones entre las medidas directas de la TVI y los índices de la atención sostenida basados en la teoría de detección de señales (sensibilidad, d' , y sesgo de respuesta, C) sobre las variables de rendimiento y comportamiento evaluadas a través de la Escala de Conducta y Rendimiento en el Aula para maestros (ECRA).

Tabla 19. Análisis de la relación entre las medidas de la TVI y las variables de rendimiento y comportamiento en el aula (ECRA).

La puntuaciones directas de la TVI presentan muy pocas correlaciones que superen el .30, y en todo caso sólo cabe destacar las de la puntuación de aciertos que está por encima del .20 en tres variables: rendimiento escolar, problemas de trabajo escolar y problemas de inatención. Por contra, las correlaciones del índice de sensibilidad (d') son más elevadas, e incluso superan el .30 en bastantes ocasiones. La mayor sensibilidad, esto es, la mayor capacidad de mantener la atención corre-

Tabla 19. Análisis de la relación entre las medidas de la TTV y las variables de rendimiento y comportamiento en el aula (ECRA)

	R.Acad.	Prob.Escolar	Inatención	Impulsividad	Prob.Social
1° (n=177)					
d'	,30	-,27	-,26	-,12	-,11
C	-,05	,03	,01	-,08	-,01
Aciertos	-,16	,20	,19	,07	,08
Comisiones	-,22	,22	,20	,13	,07
T. Reacción	,03	-,03	-,10	-,14	-,08
2° (n=135)					
d'	,29	-,30	-,33	-,08	-,17
C	-,05	,10	,05	-,11	-,09
Aciertos	-,22	,25	,25	,01	,07
Comisiones	-,23	,19	,23	,13	,19
T. Reacción	,04	,09	,00	-,03	-,12
3° (n=93)					
d'	,37	-,21	-,28	-,11	-,09
C	-,09	,16	,14	-,06	,01
Aciertos	-,28	,19	,23	,03	,06
Comisiones	-,31	,16	,18	,16	,11
T. Reacción	-,1	,16	,20	,10	,07
4° (n=139)					
d'	,36	-,23	-,37	-,14	-,18
C	-,27	,10	,12	-,06	,02
Aciertos	-,33	,22	,31	,10	,11
Comisiones	-,25	,21	,28	,18	,16
T. Reacción	,03	-,19	-,21	-,19	-,14
total (n=584)					
d'	,32	-,25	-,31	-,10	-,13
C	-,09	,09	,07	-,08	-,02
Aciertos	,24	-,21	-,20	-,05	-,13
Comisiones	-,18	,14	,16	,12	,13
T. Reacción	,04	,02	-,04	-,04	-,03

Nota: se resaltan aquellas correlaciones iguales o superiores al .30 (dado el tamaño de la muestra es posible encontrar correlaciones significativas en valores cercanos a .10 pero en la realidad se consideran muy bajas).

laciona positivamente con el rendimiento escolar y negativamente con los problemas de trabajo escolar (aunque en la muestra total sólo llegue al $-0,25$) y de inatención. En las otras dos variables, problemas de impulsividad de interacción social, su implicación no parece relevante. Finalmente el sesgo o criterio de respuesta (C) muestra correlaciones despreciables con todas las variables de la escala. Por ello en los análisis comparativos posteriores prácticamente hemos dejado de lado este índice. Además, tampoco la interacción entre d' y C parece permitir agrupar sujetos para realizar comparaciones a posteriori. Nos hemos planteado la posibilidad de trabajar esta interacción a partir de las siguientes condiciones: en primer lugar hemos agrupado a toda la muestra, aunque respetando las puntuaciones de cada sujeto en su curso, y hemos definido la presencia de un sesgo arriesgado a partir de una puntuación de C menor de -1 desviación estándar, y hemos definido un sesgo conservador a partir de una puntuación de C superior a $+1$ desviación estándar. Y, en segundo lugar, hemos hecho lo mismo en función de d' : definir una baja sensibilidad a partir de una puntuación inferior a -1 desviación estándar de d' y una alta sensibilidad a partir de una puntuación superior a $+1$, siendo la normalidad las puntuaciones intermedias.

Con estas premisas la distribución de sujetos con sesgo arriesgado o conservador en función de su d' es la que muestra la tabla 20.

Tabla 20. Distribución de sujetos de sesgo arriesgado o conservador (criterio C) en función de su índice de sensibilidad (d')

	sensibilidad baja	normal	sensibilidad alta	total
sesgo arriesgado	30	52	2	84
Porcentaje	29,7%	13,4%	2,2%	14,5%
sesgo conservador	17	68	—	85
Porcentaje	16,8%	17,5%	—	14,6%
sin sesgo definido	54	272	89	415
Porcentaje	53,5%	69,2%	97,8%	70,9%
total	101	392	91	584
Porcentaje	100,0%	100,0%	100%	100,0%

Como vemos la inmensa mayoría de sujetos no presentan ningún tipo de sesgo definido, y sólo el grupo de sensibilidad baja tiene porcentajes relativamente significativos de sujetos con sesgo arriesgado o conservador. Pero no se puede establecer una comparación entre éstos y los sujetos de sensibilidad alta porque entre ellos prácticamente no hay sujetos con sesgo definido (sólo dos de sesgo arriesgado). Por tanto, nos vamos a centrar en el índice de sensibilidad para analizar la influencia de la atención sostenida en el ámbito escolar.

En la tabla 21 se han establecido dos sistemas de clasificación de sujetos en función de su capacidad de atención sostenida: la clasificación *dicotómica* es la ya utilizada en la tabla 18 para analizar la interac-

Tabla 21. Comparación de los grupos clasificados en función de su capacidad de atención sostenida (índice d') en las variables de rendimiento y comportamiento en el aula (ECRA).

	n	R.Acad.		Prob.Escolar		Inatención		Impulsividad		Prob.Social	
		M	DE	M	DE	M	DE	M	DE	M	DE
Clas. dicotómica											
Baja sens. (1)	98	9,12	2,65	9,88	3,66	19,12	7,30	11,49	4,88	7,82	3,79
Normal (2)	341	10,15	2,73	8,86	3,69	16,20	6,90	11,62	5,10	7,67	3,60
Alta sens.(3)	76	11,36	2,34	7,41	3,27	12,88	6,37	10,08	4,71	6,74	3,30
F		14,86** (1-2-3)		9,96** (1,2-3)		17,65** (1-2-3)		3,03		2,43	
Clas. por grupos											
Muy Baja sens.(1)	46	8,37	2,86	10,52	3,61	20,04	7,77	11,28	5,34	8,09	4,18
Baja sens.(2)	52	9,80	2,26	9,31	3,64	18,31	6,84	11,67	4,47	7,58	3,43
Normal(3)	341	10,15	2,73	8,86	3,69	16,20	6,90	11,62	5,10	7,67	3,60
Alta sens.(4)	46	11,02	2,31	7,35	3,23	13,46	6,37	10,15	4,74	6,63	3,32
Muy Alta sens.(5)	30	11,90	2,34	7,50	3,37	12,03	6,37	9,97	4,76	6,90	3,33
F		9,79** (1,2-2, 3,4-4,5)		5,68** (1,2, 3-2,3,4,5)		9,41** (1,2,3, 2,3,4,5)		1,55		1,36	

* $p < .05$, ** $p < .01$

Clasificación dicotómica. Baja sensibilidad: sujetos con $d' < -1$. Alta sensibilidad: sujetos con $d' > +1$.

Clasificación por grupos. Muy Baja sensibilidad: sujetos con $d' < -1,5$. Baja sens.: sujetos entre $-1,5$ y -1 . Alta sens.: sujetos entre $+1$ y $+1,5$. Muy Alta sens.: sujetos con $d' > +1,5$.

ción entre el índice de sensibilidad y el sesgo de respuesta, sólo que esta vez en el análisis de la varianza incluimos también el grupo de sujetos «normales», es decir, aquéllos cuya puntuación en d' estaba entre $+1$ y -1 desviación estándar. En la clasificación por grupos proponemos una mayor discriminación en la clasificación de sujetos a través de distribuir grupos alejados 1 ó 1,5 +/- desviaciones estándar. No pudimos realizar las agrupaciones a partir de 2 desviaciones estándar porque recordemos que la muestra se halla «depurada» de sujetos extremos y muy pocos ocupan esas posiciones tan alejadas en la distribución de la curva normal.

En el primer sistema de clasificación propuesto, más genérico, aparecen diferencias estadísticamente significativas en tres de las cinco subescalas de la ECRA. En rendimiento académico los tres grupos se diferencian entre sí, siendo el de mejor puntuación el de alta sensibilidad y el de peor puntuación el de baja sensibilidad. En los problemas de trabajo escolar se mantiene la misma tendencia, solamente que esta vez baja sensibilidad y grupo normal no llegan a diferenciarse. Por último, En la subescala de inatención es donde las diferencias son más evidentes con mayores problemas para el grupo de baja sensibilidad y una mejor puntuación para el de alta sensibilidad.

En el sistema de clasificación por grupos, más específico, las diferencias siguen la misma tendencia apuntada ya anteriormente. De nuevo sólo las tres primeras variables resultan estadísticamente significativas. En rendimiento académico los grupos de muy baja y baja sensibilidad por un lado y los de alta y muy alta sensibilidad por otro son los extremos en la dirección esperada, pero dentro de cada uno de estos dos polos no hay diferencias estadísticamente significativas. Prácticamente ocurre lo mismo en problemas de trabajo escolar y en inatención, sólo que en estos casos incluso a veces sólo queda como grupo claramente diferenciado de todos los demás el de muy baja sensibilidad. En cualquier caso, parece evidente que a mayor capacidad de atención sostenida mejores puntuaciones en variables académicas y menos problemas de inatención detectados por el maestro.

7 CONCLUSIONES Y PERSPECTIVA DE FUTURO

7.1 Principales conclusiones en relación a los objetivos del trabajo

El objetivo principal del trabajo, es decir, la elaboración de pruebas objetivas por ordenador para evaluar la reflexividad-impulsividad y la atención sostenida, más allá de las pertinentes consideraciones que enseguida abordaremos, consideramos que se ha cumplido en general satisfactoriamente. Para desarrollar las dos pruebas que aquí hemos presentado se ha trabajado intensamente durante los últimos tres años, con casi 2.000 niños evaluados, y a la vez lo hemos compaginado con la divulgación de los resultados en diversos foros científicos, cursos y entre investigadores con muy buena acogida. Hasta el punto que la editorial Albor-Cohs ya ha publicado una versión profesional de la TEF-IMAT, llamada Escala Magallanes de Impulsividad Computarizada (EMIC, Servera y Llabres, 2000) para su uso en el ámbito infantil, y estamos en trámites de iniciar la publicación de la TVI-IMAT. Por otra parte, en sus versiones de investigación, ambas pruebas nos han sido solicitadas para participar en trabajos de investigación en nuestro propio país (por ejemplo, el grupo que dirige en la Universidad de Cádiz, el profesor Navarro Guzmán) y en Venezuela, concretamente el grupo de trabajo de la doctora Cecilia Montiel. Pero dejemos de lado estas consideraciones generales y entremos en detalle en el análisis de los objetivos propuestos.

En el caso de la TEF los resultados que apoyan su utilización, y la posterior versión más profesionalizada que hemos realizado, se basan, en primer lugar, en una correlación negativa relativamente elevada entre errores y latencias, que es el fundamento para aplicar con fiabilidad cualquier sistema de evaluación. Solamente en primer curso esta correlación ha sido algo baja, sobre el $-.51$, tal vez porque la prueba resulta algo difícil entre los 6-7 años, pero en el conjunto de la muestra ha subido hasta el $-.67$, y en general se relaciona con valores cercanos o superiores al $-.60$. Aunque el MFF20 presenta valores similares, las ventajas de poder utilizar una versión informatizada como la nuestra saltan a la

vista. En segundo lugar, los índices de fiabilidad por consistencia interna y por el procedimiento test-retest (más costoso y por ello con una muestra más reducida de sujetos) han dado resultados positivos: el coeficiente alfa de Cronbach no baja del .77 en los errores y del .88 en las latencias a través de los cuatro cursos evaluados, mientras en el test-retest los errores están alrededor del .65 y aunque en las latencias se registra más variabilidad en función del curso (fluctúa entre .55 y .97) también cabe concluir que es aceptable. Los resultados son algo inferiores que los presentados por el MFF20, pero aún a la espera de conseguir muestras más amplias para el test-retest, son, como decíamos, razonablemente aceptables. Además, aunque son datos que no hemos podido incluir en este trabajo, tenemos ya referencias de la correlación entre la TEF y el propio MFF20 que también son favorables a la validez convergente entre las pruebas. En tercer lugar, como deseábamos, la TEF presenta un efecto de maduración que hace que con el aumento de la edad de las muestras se tienda hacia la reflexividad, sin que se registren diferencias por sexos (excepto en primero con ligera superioridad de las niñas sobre los niños). El control estricto de sujetos extremos y *outliers* es el que nos ha permitido en la EMIC, a la que antes aludíamos, realizar una baremación de las puntuaciones por edades. En cuarto, y último lugar, aunque dejaremos los comentarios específicos para más adelante, también ha sido positivo la valoración del impacto de las medidas y los índices de la TEF especialmente sobre el rendimiento académico y sobre algunos aspectos comportamentales del niño.

En cuanto a la TVI también cabe hablar de resultados bastante satisfactorios. Aunque persisten dudas sobre algunos de los parámetros que la caracterizan, especialmente la duración que para algunos puede ser algo reducida (el CPT-II de Conners dura aproximadamente 14 minutos), y tampoco tenemos datos sobre aspectos de «*menoscabo atencional*», la verdad es que ha mostrado capacidad de evaluación y discriminación de los niños. En primer lugar, y aunque en este caso se trate de medidas independientes, los aciertos y los errores de comisión presentan, en la muestra total, una correlación de -.47 que, aunque variable entre cursos, representa que el niño con mejor capacidad para

detectar *targets* tiende a cometer menos errores de comisión, es decir, de falta de control motor. En segundo lugar los datos de fiabilidad que hemos podido obtener por el procedimiento test-retest también son bastante favorables. Teniendo en cuenta que la atención es un mecanismo muy variable (y más en niños), sujeto a cambios por factores orgánicos y ambientales múltiples, el hecho de que los aciertos presenten índices de fiabilidad entre .69 y .79 y que las comisiones (en teoría más ligadas a factores motivacionales, de adopción de un sesgo de respuesta, y por tanto, aún si cabe más fluctuables) estén entre .56 y .88 son datos muy positivos. De nuevo la tendencia es que en los primeros cursos los resultados sean algo inferiores probablemente porque la prueba es más exigente a esas edades. En tercer lugar también considerábamos muy positivo que se detectase un efecto madurativo sobre la TVI, y así ha sido: los niños mayores consiguen más aciertos, reducen significativamente sus errores de comisión y, aunque en menor grado, acaban siendo más rápidos en su tiempo de reacción. En términos de índices combinados, la sensibilidad o capacidad atencional aumenta claramente con la edad y el sesgo de respuesta (aunque en general no es en ningún curso especialmente destacado) también tiende a desaparecer. No hay diferencias por sexo, excepto en la propia variable del «*tiempo de reacción*» donde de manera sistemática en todos los cursos (y podríamos añadir aquí que en todos los estudios pilotos anteriores) los niños siempre presentan medias significativamente más rápidas que las niñas. Por último, en cuarto lugar, al igual que ya comentábamos en el caso de la TEF, hay suficientes indicios de que la atención sostenida tiene implicaciones importantes al menos en variables de rendimiento y aprendizaje.

Antes de entrar en estas implicaciones, nos gustaría destacar que tanto la TEF como la TVI facilitan un aspecto crucial de nuestro trabajo, que también se recogía en nuestros dos primeros objetivos generales: el cálculo de los índices combinados tanto de reflexividad-impulsividad como de atención sostenida. Al registrarse las medidas directas de cada evaluación en el ordenador esta labor, a menudo engorrosa y a veces inexacta con el cálculo manual, adquiere elementos de mayor objetivi-

dad y facilita la labor tanto del investigador en este caso, como en el futuro del profesional.

Entrando ya en lo que son los objetivos generales 3 y 4, es decir, las implicaciones de cada dimensión en variables medidas en el aula, hay que empezar destacando que las dos coinciden en presentar sus correlaciones más elevadas, por este orden, sobre las variables de «rendimiento académico», «problemas de trabajo escolar» e «inatención», es decir, y aunque después lo justificaremos más claramente, variables que podemos considerar del ámbito cognitivo en general, mientras en «impulsividad» y «problemas de interacción social», sin duda variables más de orden *comportamental*, su incidencia es mucho menor.

En el caso concreto de la TEF se observan un grado razonable de variación por curso, sin que necesariamente se pueda destacar una tendencia clara, y al final nos encontramos con una superioridad del índice de estilo (PI) sobre el de eficacia (PE). Ello quiere decir que, en términos generales, presentar un estilo tendente a la impulsividad supone, en opinión de los maestros, mayores probabilidades de tener un rendimiento académico bajo y más problemas de trabajo escolar. La siguiente correlación es la de la subescala de inatención, pero mientras las anteriores en la muestra total superan el .30, ésta baja al .24. Dado que la PI se compone de las puntuaciones directas de errores y latencias, nuestros datos confirman la conclusión ampliamente destacada en trabajos previos que la contribución de los errores es siempre mayor que la de latencias. Se trata, en términos generales, de unos 11-12 puntos en el conjunto de las correlaciones, sin embargo, no pensamos que este hecho deba desembocar en la idea de la absoluta inocuidad de la puntuación de latencias: en primer lugar porque, aún siendo más bajas, las correlaciones de las latencias son en la dirección esperada. Y, en segundo lugar, porque es la PI y no los errores solos la variable que detecta más implicaciones en el aula. Si comparamos la PI con los errores sobre las variables de la ECRA y obtenemos el coeficiente de determinación de sus correlaciones en la muestra total podemos concluir que aproximadamente un 12% del rendimiento académico, un 9% de los problemas de trabajo escolar y un 6% de la inatención observada en el aula por los maestros es explicable

a través del índice PI. En el caso de los errores su contribución es respectivamente del 9%, 9% y 4%, es decir, algo inferior. En el caso de la PE, el índice de eficacia, estos mismos valores bajan hasta el 4% en el rendimiento y en los problemas del trabajo escolar (siendo despreciable su aportación a la subescala de inatención).

En conclusión, por lo que respecta a la reflexividad-impulsividad, se trata de una dimensión claramente influenciada por un componente de capacidad, que tal vez desautoriza su conceptualización en términos ortodoxos de «*estilo cognitivo*», pero que la refuerza en términos de estilo de procesamiento o de aprendizaje. A través de la TEF (como antes ya se había observado sobre las MFF) la concepción estilística, entendida como un enfoque de la tarea basado en la reflexión (el despliegue de estrategias de afrontamiento y resolución del problema), sigue siendo la que hace más aportaciones a las variables académicas. En todo caso, aún si tomásemos a los errores como principal variable predictiva, hay que aceptar que la mejor manera de cometer pocos errores es adoptando precisamente la actitud reflexiva. Ha constituido una relativa sorpresa que la subescala de inatención, aunque en menor grado, se haya incorporado a las correlaciones relativamente elevadas con la PI. Nuestra explicación es que los ítems de esta subescala (no acabar las actividades, no concentrarse, soñar despierto, no seguir instrucciones, distraerse, etc.), como casi todos los de «*tipo atencional*» de cualquier otro cuestionario o escala, en realidad no sólo hacen referencia a esta dimensión, sino también a capacidades cognitivas genéricas. Estas capacidades incluyen aspectos relacionados con los procesos de resolución de problemas y de autocontrol. Al tocar este ámbito la dimensión R-I entra en juego, y rápidamente, aunque sea de soslayo, ya aparecen los datos que relacionan estas capacidades con el estilo. Por otra parte, y ello sí que no constituye ninguna sorpresa, en las subescalas de «*impulsividad*» y «*problemas de interacción social*», con ítems más claramente relacionados con el comportamiento social (hablar demasiado, estar inquieto, no poder aguardar turno, ser algo agresivo, etc.), la R-I prácticamente ya no presenta ninguna influencia. Con todo se confirma la orientación cognitiva restringida, centrada en aspectos del procesa-

miento de información que se relacionan con las habilidades y estrategias de aprendizaje escolar, de la dimensión y de nuevo se descarta su influencia genérica en otros ámbitos.

En el caso de la TVI la tendencia general es similar a la TEF sólo que la magnitud de las correlaciones es inferior prácticamente en todas las variables (aunque muy cercanas o algo por encima del .30). Las principales correlaciones también se concentran en la variable «*rendimiento académico*», pero en este caso la subescala de «*inatención*» prácticamente iguala o supera a la de «*problemas de trabajo escolar*». Las puntuaciones directas de aciertos y errores de comisión presentan unos niveles de correlación similares, algo superiores a los aciertos, pero en general el mejor indicador de la influencia de la atención sostenida sobre la escala lo da el índice de sensibilidad (d'). La aportación del otro índice, el sesgo de respuesta (criterio C) y de la medida del tiempo de reacción (una medida muy básica para que destaque en muestras no clínicas) es prácticamente nula. El índice de sensibilidad (la resta de las puntuaciones tipificadas de aciertos y errores de comisión) es, al menos desde nuestro punto de vista, el referente directo de la capacidad de atención sostenida del niño. Como era de esperar a mayor capacidad atencional mejor rendimiento académico, menos problemas en el trabajo académico y, lo que resulta más interesante, menos conductas de «*inatención*» según la percepción de los maestros. Este último valor favorece la validez convergente de la TVI aunque, como ya hemos explicado, entendemos que en los ítems de esta subescala no sólo se hallan presentes componentes atencionales, sino también de procesos ejecutivos en general. Cerca del 10% de la variabilidad en el rendimiento académico y de los problemas de inatención en el aula se puede explicar a través del índice de sensibilidad de nuestra prueba de atención sostenida (recordemos, una tarea simple de vigilancia de cerca de siete minutos y medio de trabajo), mientras algo más del 6% de los problemas de trabajo escolar también dependen de él. El índice supera la aportación individual de los aciertos o los errores de comisión. Por último, las bajas correlaciones del sesgo de respuesta deben contemplarse con precaución; como ya hemos explicado, al depurar la muestra definitiva de suje-

tos extremos ha podido pasar que muchos con sesgo de respuesta muy acentuado hayan quedado fuera de los análisis estadísticos. En los estudios con muestras clínicas, o más específicamente para la detección de niños con dificultades de aprendizaje, no como en nuestro caso que buscábamos una aproximación genérica al tema, es posible que el sesgo de respuesta tenga mucha más incidencia de la que aquí presenta.

En conclusión, por lo que respecta a la atención sostenida, nuestra tarea de vigilancia, aún con las limitaciones teóricas que presenta, se ha mostrado con capacidad de discriminación y de predicción de variables relacionadas con el funcionamiento cognitivo y el aprendizaje del niño. Por contra, no ha mantenido influencia sobre las variables más estrictamente comportamentales. El lector no debe estar tentado de negar validez de constructo a la TVI por el hecho de ser el déficit en atención sostenida muy típico de los trastornos por hiperactividad, y de no relacionarse aquí con las subescalas de «impulsividad» y problemas de conducta en general, típicos de estos niños. Las diferencias significativas se han extraído sobre muestras *extremadamente normalizadas* (valga la expresión) y en la práctica las diferencias en puntuación entre grupos de sujetos son mínimas. Nuestra pretensión era desde el principio, en todo caso, ofrecer parámetros de baremación sobre los cuales en un futuro comparar la ejecución de los niños con sospecha de padecer trastornos por déficit de atención y/o hiperactividad, pero no en este momento de la investigación. Por tanto, nos damos por muy satisfechos que la tarea de vigilancia siga siendo capaz de mostrar repercusiones de la atención sostenida sobre el ámbito del aprendizaje y de la actuación cognitiva en general, una vez se han eliminado casi 185 sujetos de la muestra inicial por múltiples razones, pero entre ellas probablemente por padecer trastornos por déficit de atención y de comportamiento en general.

Las comparaciones a través de los análisis de variancia y dos a dos han servido, de entrada, para confirmar las diferencias estadísticamente significativas sobre las variables de rendimiento y comportamiento que se intuían a través de las correlaciones. Así los niños reflexivos se diferencian, en general, de los impulsivos (y por orden en función del valor

del estadístico f) en rendimiento académico, problemas de trabajo escolar, inatención, impulsividad y problemas de interacción social. Mientras los de alta capacidad atencional lo hacen a su vez de los de baja capacidad con la misma tendencia, pero sin llegar a la significación en las dos últimas variables. En todo caso, hay que matizar estos resultados en función del sistema de clasificación.

En el caso de la reflexividad-impulsividad teníamos mucho interés en comprobar el funcionamiento de nuestro sistema de clasificación, al que hemos denominado IMAT, porque detectamos problemas conceptuales en el funcionamiento de los otros dos más tradicionales: el de cuadrantes y el integrado. En realidad, a primera vista, los tres sistemas detectan diferencias significativas en casi todas las variables en parecidos términos. Pero un análisis más detenido permite destacar las ventajas del sistema IMAT. El sistema integrado de Salkind y Wright intentó superar las limitaciones de los cuadrantes, a través del uso de la PI y la PE, como hemos explicado sobradamente en este trabajo. Pero en realidad ni ofrece puntos de corte (es decir, si la media cero de la PI diferencia a impulsivos y reflexivos, y se opta por este punto, se incorporan al mismo grupo niños con tendencia similar pero de magnitud muy diferente), ni cuando se adoptan (por ejemplo, sujetos impulsivos superior a +1 desviación estándar, y reflexivos inferior a -1 desviación estándar) se puede garantizar la homogeneidad de estos grupos: ¿qué impide que en el grupo de impulsivos no convivan impulsivos puros (con un nivel de eficacia, medido por la PE, normal), rápidos-exactos (impulsivos con una elevada eficacia) y rápidos-inexactos (impulsivos que además presentan un pésimo nivel de eficacia sobre la prueba)? El sistema IMAT evita esta posibilidad al seleccionar, en primer lugar, el grupo mayoritario de sujetos de estilo flexible y eficacia normal (+/- 1 desviación estándar) y, a continuación, permitir la selección de impulsivos puros, reflexivos puros, ineficaces puros, eficaces puros y los grupos en interacción: impulsivos eficaces o ineficaces y reflexivos eficaces o ineficaces. Con estas premisas el sistema IMAT nos ha permitido extraer estas conclusiones, que no se observan en los otros sistemas:

- En la clasificación basada en la PI, y en una muestra sin sujetos extremos, impulsivos y flexibles han formado un grupo único (aunque los primeros tenían peores puntuaciones) con puntuaciones inferiores en rendimiento a los reflexivos, y superiores en «*problemas de trabajo escolar*» y especialmente «*inatención*», mientras el nivel de significación se ha reducido mucho en las subescalas de «*impulsividad*» y «*problemas de interacción social*».
- En la clasificación basada en la PE la composición de los grupos ineficaz y eficaz es mucho menor, puesto que el buen comportamiento de la correlación entre errores y latencias en la TEF deja poco margen para estos grupos. Los resultados son en la misma línea que en la anterior clasificación, favorables al grupo eficaz, sólo que aquí los problemas de interacción social también asumen un cierto nivel de significación, siendo menos habituales en el grupo mencionado.
- En la clasificación en interacción hemos podido comprobar que, aunque como es lógico el número de sujetos es todavía más reducido, si pueden influir decisivamente si no se eliminan de la clasificación en estilo, como hemos hecho nosotros. De nuevo hay que apelar a que si en una muestra depurada, donde la mayoría de sujetos extremos han quedado fuera, los que han podido mantenerse siguen presentando las peores (impulsivos ineficaces) y mejores (impulsivos eficaces) puntuaciones en las variables académicas. En otras investigaciones donde no se ha tomado esta precaución estos sujetos pueden alterar gravemente las conclusiones que se hagan en nombre exclusivamente de un estilo impulsivo o reflexivo.

En cuanto a los sistemas de clasificación basados en las puntuaciones de la tarea de vigilancia el trabajo ha resultado más sencillo, al no ser necesario el hacer depender el índice de sensibilidad del sesgo de respuesta. Sólo con el índice de sensibilidad nos hemos limitado a comprobar si una clasificación *empírica* y dicotómica (con sujetos de alta y

baja capacidad, además del grupo de normalidad), basada en +/- 1 desviación estándar del valor de d' , era suficiente o, por el contrario, era mejor trabajar con una clasificación más amplia de hasta cinco grupos, también definidos en función de desviaciones estándar. Los resultados han sido similares, pero el primer sistema, más simple, ofrece la lectura más clara: los sujetos de alta sensibilidad, sensibilidad normal y baja sensibilidad se diferencian entre sí en las variables de «rendimiento académico» e «inatención»; los dos últimos se agrupan frente al primero en «problemas del trabajo escolar» y no hay diferencias en las subescalas de «impulsividad» y «problemas de interacción social». Todo ello siempre en la dirección que ya marcaban las correlaciones: la capacidad atencional tal como la medimos nosotros es una variable relevante en el ámbito educativo.

El quinto objetivo general de nuestro trabajo cobra más relevancia en función de las conclusiones expuestas hasta el momento: si impulsividad y baja atención se relacionan con problemas académicos y de conducta inatenta en el aula casi por igual, ¿hasta qué punto en estos grupos se hallan los mismos niños? La relación que hemos encontrado entre la TEF y la TVI depende, en primer lugar del grupo de edad en que nos centramos y, en segundo lugar, de si trabajamos con las puntuaciones directas o con el índice de sensibilidad. La relación entre estilo reflexivo/impulsivo y baja/alta atención es en la dirección esperada: a mayor reflexividad más capacidad atencional. Ahora bien, se trata de una relación moderadamente elevada sólo en los cursos de tercero y cuarto (alrededor del .40) y entre la PI y d' . La puntuación de eficacia y el sesgo de respuesta presentan valores muy bajos o incluso nulos. Si observamos las puntuaciones directas nos encontramos, en la muestra total, con una correlación de los errores de la TEF de -.56 con los aciertos y de .45 con los errores de comisión de la TVI, mientras las latencias lo hacen respectivamente con .36 y -.35, es decir, de manera inferior pero significativa. Los errores de la TEF correlacionan sólo algo más que las latencias con d' (-.23 versus .18 en la muestra total) y, por su parte, los aciertos y las comisiones de la TVI lo hacen por igual sobre la PI (alrededor del .20, siendo obviamente negativa en el caso de los aciertos).

En conclusión existe una correlación moderadamente elevada entre las dimensiones reflexividad-impulsividad, medida por la TEF, y atención sostenida, medida por la TVI, que, en términos cuantitativos puede expresarse del siguiente modo (tomando como punto de referencia el sistema dicotómico para clasificar a los niños en función de su capacidad de atención o d' y la clasificación en estilo del sistema IMAT): un 78% de los niños con baja sensibilidad son también niños impulsivos, mientras sólo un 22% son reflexivos. Al revés, un 73% de los niños con alta sensibilidad son también reflexivos, mientras sólo un 27% son impulsivos. El valor del chi-cuadrado en la tabla de contingencia es del 23,39 ($p < .01$) indicando una clara dependencia entre ambas variables. Por tanto, al menos en lo que a nuestros conocimientos respecta, esta es una de las primeras ocasiones en que, a través de pruebas objetivas, se logra establecer una clara relación entre el estilo impulsivo y una medida de la capacidad atencional. Hay que reconocer que ello ha sido posible a través de índices y sistemas de clasificación relativamente complejos, lo cual deja en entredicho las múltiples investigaciones que sólo se valen de las puntuaciones directas para establecer conclusiones al respecto. Los índices son los valores realmente relevantes en este tipo de dimensiones, aunque nada impide que la investigación futura pueda mejorar su cálculo o su significación.

Para finalizar las conclusiones referentes a este objetivo general queremos recordar que hay un efecto de maduración todavía no identificado que media en la relación entre impulsividad y baja atención: es a partir de los 8 años y medio aproximadamente cuando los datos avalan más claramente su presencia. Sería más lógico, desde el punto de vista de las teorías del desarrollo cognitivo, que inatención e impulsividad compartieran más relación en las edades más tempranas, y a medida que se produce la maduración de sus mecanismos implicados se vayan diferenciando, pero curiosamente ha sido al revés. Hay varias hipótesis que pueden explicar este hecho: en primer lugar cabe la posibilidad de que las dos pruebas sean demasiado difíciles para muchos de los niños de menor edad, y las ejecuciones se hallen muy contaminadas por procesos de ensayo/error. En segundo lugar, pero relacionada con la anterior,

también es posible que los procesos y las estrategias de resolución de la prueba en principio más compleja, la TEF, cambien con la edad: en los primeros años todo puede depender en mayor grado de una capacidad y/o de un estilo de procesamiento general, mientras a partir de los 8/9 años procesos, mecanismos y operaciones cognitivas más específicas, como la atención sostenida, van modificando el modo de afrontar no sólo la TEF, sino las tareas de resolución de problemas en general. Como ya comentamos en la revisión del tema que hicimos en la primera parte del trabajo, se cree que el proceso de maduración de la atención sostenida es más lento que el de la atención selectiva, y que es precisamente sobre las edades clave que nosotros hemos detectado cuando se produce un salto cualitativo en la capacidad del niño para su uso.

Para acabar, y en relación a los cuatro objetivos específicos que nos planteábamos, aunque ya hemos comentado en mayor o menor medida bastantes aspectos que hacen referencia a ellos podemos concluir que: como esperábamos prácticamente no ha habido diferencias por sexos ni en reflexividad-impulsividad, ni en atención sostenida, aunque la medida del tiempo de reacción siempre ha sido significativamente favorable a los niños. En cambio, sí que se han detectado diferencias por cursos, aumentando la reflexividad y la capacidad de atención sostenida con la edad. En cuanto al segundo objetivo específico, aunque todos los sistemas de clasificación del estilo impulsivo pueden ser útiles para encontrar diferencias con respecto al reflexivo, nuestra propuesta (el sistema IMAT) presenta todas las ventajas que ya hemos mencionado anteriormente, además de una mayor claridad en la conceptualización de las puntuaciones directas y los índices combinados. En el caso de la atención sostenida, por un lado, cabe destacar como fundamental la aportación de al menos uno de los índices derivados de la teoría de la detección de señales, el de sensibilidad, puesto que si solo se trabaja con las puntuaciones directas de aciertos y comisiones se hace difícil interpretar correctamente el nivel de atención que cabe atribuir a la actuación del niño. Por otro lado, y aunque es un tema que requiere un mayor desarrollo, la clasificación dicotómica a partir de la sensibilidad que da d' resulta de entrada suficiente para detectar diferencias especial-

mente en variables relacionadas con el aprendizaje y las conductas inatentas de los niños en el aula.

En definitiva pensamos que hemos conseguido desarrollar dos pruebas de evaluación, ligadas a dos dimensiones cognitivas claves en el ámbito del aprendizaje, que tienen un amplio campo de aplicación tanto desde el punto de vista profesional, en el ámbito académico y clínico infantil, como en el campo de la investigación, donde sin duda todavía quedan muchos aspectos por analizar.

7.2 El futuro de la evaluación objetiva de la impulsividad y la inatención en el aula

En los aspectos de futuro de nuestra línea de trabajo hay que distinguir aspectos más de ámbito profesional, que ya casi podemos decir que son presente, de otros más de investigación a largo plazo. Entre los primeros, como ya hemos mencionado brevemente, cabe resaltar que la TEF-IMAT ya tiene una versión aplicada, la EMIC (Servera y Llabrés, 2000) que se está comercializando. Nosotros hemos utilizado la EMIC, y también la TVI, en una Unidad de Hiperactividad del Departamento de Psicología de la Universidad de las Islas Baleares que ha funcionado los últimos dos años. Aproximadamente 150 niños han sido derivados por pediatras de atención primaria, neurólogos, equipos psicopedagógicos de la Conselleria de Educación, o profesionales de los propios centros escolares con sospecha de padecer trastornos por déficit de atención con hiperactividad. Entre el amplio protocolo que utilizamos para establecer nuestra hipótesis diagnóstico están las dos pruebas aquí presentadas, y aunque es pronto para establecer conclusiones, los resultados parecen ser positivos. Hay que tener en cuenta, como ya vimos en su momento, que tanto las versiones de los tests de reflexividad-impulsividad como de atención sostenida han sido puestos en duda especialmente en su capacidad para ayudar al diagnóstico de los trastornos de hiperactividad (véase la interesante revisión de Rapport y cols., 2000). Ello realmente es un problema porque estas medidas son de las más

objetivas que disponemos para acercarnos a este complejo trastorno y, por tanto, cabe insistir en esta línea de investigación, mejorando si es necesario las pruebas hasta definir exactamente cuál es su potencial en el ámbito clínico.

En lo que respecta a las versiones profesionales, tanto la de la EMIC como la que estamos preparando de la TVI, este trabajo ha servido, en primer lugar, para conseguir establecer datos de baremación por distintos niveles de edad y sistemas de clasificación de los sujetos relativamente fiables y válidos. Sin duda estas eran las cuestiones que, hasta el momento circunscribían a la reflexividad-impulsividad y en mayor medida a la atención sostenida a la investigación básica, pero ahora les abren las puertas al trabajo clínico diario, con las reservas antes apuntadas que deberán resolver los investigadores.

Entre las cuestiones abiertas a medio y largo plazo que afectan directamente a nuestra línea de trabajo es posible distinguir entre aquellas más de tipo metodológico de aquellas más de tipo conceptual, aunque es evidente que se trata de una división artificial para facilitar la exposición de motivos.

En el ámbito metodológico nos queda, en primer lugar, y referente a la reflexividad-impulsividad conseguir más datos sobre la fiabilidad test-retest de la TEF/EMIC, que hasta el momento es algo inferior a la del MFF20, pero que ha sido poco explorada. También necesitamos información más detallada del sistema de clasificación IMAT. Nuestra propuesta requiere replicaciones para confirmar que aporta datos diferenciales al sistema de cuadrantes y al sistema integrado, es decir, que tiene mayor capacidad de discriminación, y que ella puede que no sólo se refleje en el ámbito académico sino también en el clínico (¿qué tipo de grupo es el que puede caracterizar a los niños con trastornos hiperactivos?). Por otra parte, nos gustaría conseguir datos de todo el nivel de primaria, y no sólo de los cuatro primeros cursos que es de lo que hasta ahora disponemos. Ello liga con la necesidad no sólo de estudios comparativos como el nuestro, sino también longitudinales que, aunque sabedores de la dificultad que implican, en el ámbito infantil y en dimensiones tan ligadas a aspectos madurativos se hacen prácticamente

imprescindibles. Por último, en el ámbito de las implicaciones, habría que utilizar procedimientos complementarios a las escalas de valoración de los profesores: sistemas de evaluación del rendimiento basados en el curriculum, pruebas estandarizadas por materias, registros de observación en ambiente natural y en el laboratorio, pruebas psicofisiológicas, etc.

Prácticamente todo lo dicho anteriormente es aplicable también a las pruebas de evaluación de la atención. En este caso, pero, podemos destacar la necesidad de evaluar no sólo la función de mantenimiento atencional, sino que se debería realizar un proceso similar al menos para la selectividad y probablemente para el cambio atencional. Existen bastantes pruebas experimentales de estos dos mecanismos que también podrían adaptarse al ámbito aplicado. En lo que respecta específicamente a la TVI hay diversas labores de investigación que requieren atención: el estudio de si se da o no menoscabo atencional (estamos preparando una versión con medidas intermedias), la posibilidad de tener que aumentar la duración de la prueba, la influencia que puede tener incluir elementos complementarios, como tareas *go- no go* (obligan a demostrar control motor al establecer estímulos bajo los cuáles no se debe responder a los *targets*) o procedimientos de ejecución continua (es decir, de responder ante cada estímulo si es o no *target*, y no limitarse a la vigilancia que hemos exigido hasta el momento), etc. También el sistema de clasificación de sujetos en función de su sensibilidad y sesgo de respuesta lo hemos desarrollado poco hasta el momento, de hecho sólo hemos trabajado con el índice de sensibilidad. Son necesarios más estudios en donde se comparen las implicaciones académicas y comportamentales de los niños con alta/baja sensibilidad que además presentan sesgo conservador o arriesgado, o no presenten sesgo como ha sido nuestro caso. Sin duda aquí hay una labor compleja, que todavía puede complicarse más si además incluimos, como se ha demostrado aquí, la posibilidad de que la dimensión atencional interactúe en el estilo reflexividad-impulsividad.

En relación a las cuestiones conceptuales sería arduo entrar en detalle en todos los aspectos polémicos o desconocidos que hemos ido

apuntando en la parte de revisión de la reflexividad-impulsividad y la atención sostenida. En realidad nuestro trabajo no aporta soluciones concretas, puesto que tampoco era nuestro objetivo, sino más bien abrir una línea que nos pudiera llevar a estas consideraciones. En cualquier caso hay algunos aspectos de primera mano que han llamado nuestra atención.

En primer lugar nos hemos referido reiteradamente a la posibilidad de que la TEF y la TVI sean algo difíciles para los niños de primer y segundo curso. Hay bastantes datos indicativos de esta posibilidad que, en la práctica, puede suponer que intentemos realizar conclusiones sobre tipos de funcionamiento cognitivo cuando en realidad ha habido un exceso ensayo/error en la actuación del niño. Hemos intentado controlar ésto con criterios estrictos para eliminar aquellos sujetos que, dentro de su grupo edad, se mostraban extremos en alguna de las puntuaciones, pero somos conscientes que tal vez ello no ha sido suficiente. En el caso de las teorías sobre la reflexividad-impulsividad, y probablemente también lo podríamos hacer extensible a la atención sostenida, Pedro Solís-Cámara (1987) ha defendido un modelo probabilístico azaroso por el cual se indica que antes de clasificar a los sujetos en función de su estilo impulsivo es necesario eliminar a aquellos sujetos que en realidad han afrontado la prueba (sea por falta de capacidad o por problemas motivacionales) por ensayo/error. El autor ha trabajado puntos de corte de las puntuaciones de error sobre versiones de los MFF a partir de los cuáles diferenciar «azarosos» de «cognitivos», y luego sólo actuar sobre los segundos. Sin duda de esta manera, al menos en teoría, se refuerza el componente estilístico de la dimensión, y se minimiza el de capacidad. Hemos tenido ya contactos con el grupo de trabajo del doctor Solís-Cámara en México y es posible que pronto sometamos a la TEF a este tipo de análisis.

En cualquier caso, y con independencia de los resultados más o menos favorables que puedan darse, hemos dejado clara nuestra postura de que no pensamos que sea una *catástrofe* que la R-I quede conceptualizada como estrategia de aprendizaje en lugar de estilo cognitivo puro. Hay datos suficientes para que, aunque no pueda librarse de cier-

ta influencia del CI o del factor G de inteligencia (como no lo hace ninguna tarea con un mínimo de capacidad de discriminación), pueda mantener una independencia que asienta su solidez sobre las implicaciones en el rendimiento y las dificultades de aprendizaje del niño.

Algo similar puede darse en la TVI y la atención sostenida en general. Puede que nuestra tarea de vigilancia no pueda considerarse, en sentido estricto, una medida pura de atención sostenida, pero sí refleja una capacidad claramente relacionada con esta dimensión. Sin duda hay que aportar más datos que confirmen la tendencia aquí apuntada, en el sentido que nuestra medida es capaz de detectar diferencias en el ámbito educativo y, sobre todo, mantiene un cierto grado de relación con la escala de valoración de conductas de inatención del maestro. La naturaleza de esta medida, como hemos dicho, puede que deba analizarse más, pero sin duda entenemos nosotros que siempre se moverá en el campo atencional. La aportación de la teoría de detección de señales ha resultado crucial para nuestros intereses. De hecho en la primera parte de nuestra investigación trabajábamos con las puntuaciones directas de los aciertos y las comisiones y siempre nos encontrábamos con resultados confusos y de difícil interpretación, y aunque intentamos combinarlas no conseguíamos aclarar mucho el panorama. Los índices de la teoría de detección de señales se constituyen en la principal apuesta cara al futuro. Ahora bien, también queremos enfatizar que la tipificación de las puntuaciones directas que lleva al cálculo de estos índices debe realizarse siempre por grupos de edad y nivel de desarrollo muy bien definidos, puesto que al tratarse de una dimensión fácilmente influible por aspectos de maduración las comparaciones entre sujetos de distinta edad pueden verse muy alteradas. Hasta el momento hemos controlado la edad y los sujetos extremos, pero puede que en el futuro haya que incorporar también medidas de desarrollo evolutivo.

El siguiente paso, tanto metodológico como conceptual, y sin duda el de mayor interés para padres y educadores es el de tratamiento. Los trabajos puramente de evaluación, como el nuestro, sólo suelen despertar interés entre algunos investigadores, porque padres y educadores se centran casi siempre en soluciones *prácticas* para los problemas de

sus niño, olvidando que para llegar a este punto en buenas condiciones hay que dar antes bastantes pasos. Nosotros hemos trabajado la modificación de la impulsividad cognitiva, y versiones anteriores tanto de la TEF como de la TVI han demostrado ser susceptibles al tratamiento cognitivo-conductual y/o farmacológico. No es un problema encontrar programas de tratamiento de la impulsividad y el déficit de atención, incluso diríamos que hay demasiada *oferta* en el mercado. El problema es que muchas veces cuando estos programas se han sometido a validación empírica muestran muchas debilidades: a veces sólo consiguen mejoras en pruebas de laboratorio, otras son muy puntuales, inestables, poco generalizadas, etc. Sin duda hay múltiples factores que inciden sobre estos resultados, pero a nuestro juicio uno de los primeros es el no haber detectado y definido claramente el tipo de problema que presenta el niño.

Se han constituido muchos grupos de niños impulsivos destinados a intervención, como es nuestro propio caso sin ir más lejos, sin tener seguridad ni del tipo de impulsividad ni del grado de impulsividad de estos niños. No puede ser lo mismo un niño extremadamente impulsivo con pésimo nivel de eficacia que otro con impulsividad moderada sin este componente, aunque en los dos se aprecien problemas de rendimiento académico. Tampoco puede ser lo mismo un niño con conductas de inatención en el aula, sin déficit sobre tareas de vigilancia, que uno que presente problemas en los dos procedimientos. Si algo se aprecia en nuestro trabajo es que existen múltiples tipos de niños que pueden acarrear distintos problemas en ambas dimensiones. Y eso sin tener en cuenta aspectos de interacción (impulsivo e inatento) y comorbilidad: presencia de hiperactividad, dislexia, trastorno de conducta, etc., que aún puede dificultar más la labor. En nuestro caso deseamos lo antes posible ponernos en marcha en el ámbito del tratamiento, pero deseamos hacerlo a partir del sistema más sólido de evaluación disponible. Ya sabíamos que existen múltiples escalas para padres y educadores, entrevistas clínicas, sistemas de observación conductual, pero pensábamos que la evaluación objetiva, de laboratorio, era la más débil, y por eso a ella hemos dedica-

do nuestros esfuerzos. Aunque, como decíamos, hay otros muchos factores implicados que pueden afectar a los resultados del tratamiento, si llegamos a definir, en forma de árbol de decisión flexible, procedimientos más o menos adecuados para tipos o grados de impulsividad y/o disfunción atencional es posible que lleguemos a desarrollar una opción tecnológica todavía hoy en día sumida en demasiada confusión.

Para acabar, y al hilo del último comentario, queremos reiterar que desde el principio nuestra intención era defender la opción de la medida objetiva, a través de tareas de laboratorio, de dimensiones cognitivas que a nuestro juicio escapan de los procedimientos de evaluación subjetiva y de los registros de observación de conducta manifiesta. Los datos, al final del trabajo, nos reafirman en esta posición y la fortalecen. Somos conscientes que sigue habiendo muchos autores críticos con esta postura (véase la ya comentada revisión de Rapport y cols., 2000), y en parte comprendemos perfectamente las limitaciones y los problemas para adaptar pruebas que sirven casi exclusivamente para el desarrollo de teorías psicológicas, a veces de base muy inferencial, al ámbito del funcionamiento diario del niño. Es aquello que con frecuencia nos han comentado con cierta sorna en cursos y reuniones de que *«todo maestro sabe si su alumno tiene o no buena capacidad atencional, sin necesidad de pasar por un ordenador»*. Más allá de que en este tipo de posturas a veces se esconden creencias filosóficas o educativas alérgicas al modelo científico-experimental, la verdad es habría que recordarles que existe la posibilidad de confundir *«saber»* con *«intuir»* y que probablemente de intuiciones, como de buenas intenciones, está el *«infierno»* (en este caso las aulas de recuperación o de educación especial) empedrado.

En cualquier caso, y sin intención de entrar en ningún tipo de polémica, simplemente queremos dejar constancia que existe en nuestro país una línea de trabajo, en la que nos inscribimos nosotros, en la que bastantes investigadores creen en una aproximación científica y aplicada a la realidad de los problemas de aprendizaje y de conducta de nuestros alumnos. Que existe un ingente trabajo en psicología básica y

experimental que hasta el momento ha pasado en gran parte desapercibido para la actividad profesional del educador y del psicólogo infantil, pero que es posible reactivar. La medida lo más exacta posible de nuestros problemas es la mejor garantía para el desarrollo de soluciones estables y generalizables. Ello pasa, en nuestra opinión, por la construcción de puentes sólidos y amplios de miras entre investigadores básicos y profesionales, y en esa labor nos estamos aplicando nosotros.

REFERENCIAS BIBLIOGRÁFICAS

- Achenback, T.M. y Edelbrock, C.S. (1983). *Manual for the Teacher Report Form and the Child Behavior Profile*. Burlington: University of Vermont, Dept. of Psychiatry.
- APA (1995). *DSM-IV. Manual diagnóstico y estadístico de los trastornos mentales*. Barcelona: Masson.
- Ault, R.L. (1973). Problem-solving strategies of reflexive, impulsive, fast-accurate, and slow-inaccurate children. *Child Development*, 44, 259-266.
- Ault, R.L., Mitchell, C. y Hartman, D. (1976). Some methodological problem in reflection-impulsivity research. *Child Development*, 47, 227-231.
- Banta, T.J. (1970). Test for the evaluation of early childhood education: The Cincinnati Autonomy Test Battery (CATB). En J. Hellmut (Ed.), *Cognitive studies* (vol. 1). Nueva York: Brunner-Mazel.
- Baños, R. y Belloch, A. (1995). Psicopatología de la atención. En A. Belloch, B. Sandín y F. Ramos (Eds.), *Manual de psicopatología* (vol. 1, pp. 165-186). Valencia: McGraw-Hill.
- Barkley, R.A. (1990). *Attention deficit hyperactivity disorder*. Nueva York: Guilford.
- Barkley, R.A. (1997). *ADHD and the Nature of Self-Control*. Nueva York: Guilford.
- Baron, J. (1985). *Rationality and Intelligence*. Cambridge, England: Cambridge University Press.
- Barret, D.E. (1977). Reflection-Impulsivity as a Predictor of Children's Academic Achievement. *Child Development*, 47, 1443-1447.
- Bentler, P. y McClain, J. (1976). A multitrait-multimethod analysis of reflection-impulsivity. *Child Development*, 47, 218-226.
- Bermejo, V. (1987). Procesos atencionales y aplicaciones educativas. *Psiquis*, VIII, 46-53.
- 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 (5), 611-632.
- Block, J., Block, J.H. y Harrington, D.M. (1975). Comment on the Kagan-Messer reply. *Developmental Psychology*, 11, 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. *Developmental Psychology*, 22 (6), 820-830.

- Bornas, X. (1994). *La autonomía personal en la infancia: Estrategias cognitivas y pautas para su desarrollo*. Madrid: Siglo XXI.
- Bornas, X. y Servera, M. (1992). Cognitive training programs to reduce impulsivity-related achievement problems: The need of in-classroom interventions. *Learning and Instruction, 2*, 89-100.
- Bornas, X., Servera, M., & Montaña, J. J. (1998). La medición de la impulsividad en preescolares: Análisis psicométrico de la escala KRISP. *Psicothema, 10*(3), 597-608.
- Bornas, X., Servera, M., Serra, F y Escudero, J.T. (1990). El tratamiento de la impulsividad infantil: Autoinstrucciones versus solución de problemas. *Estudios de Psicología, 43-44*, 61-71.
- Branningan, G.C., Ash, T. y Margolis, J. (1980). Impulsivity-reflexivity and children's intellectual performance. *Journal of Personality Assessment, 44* (19), 41-43.
- Broadbent, D.E. (1958). *Percepción y comunicación*. Madrid: Debate (trad. cast., 1983).
- Brown, R.T. y Wynne, M. (1982). Correlates of teacher ratings, sustained attention, and impulsivity in hyperactive and normal boys. *Journal of Clinical Child Psychology, 11*, 262-267.
- Brown, R.T. y Wynne, M.E. (1984). Attentional characteristics and teachers ratings in hyperactive reading disabled and normal boys. *Journal of Clinical Child Psychology, 13*(1), 38-43.
- Buela-Casal, G., Carretero-Dios, H. y De los Santos-Roig, M. (2001a). Análisis del constructo reflexividad-impulsividad: Del Matching Familiar Figures Test (MFFT) al MFFT20. *Análisis y Modificación de Conducta, 27*, 29-70.
- Buela-Casal, G., Carretero-Dios, H. y De los Santos-Roig, M. (en prensa). Consistencia longitudinal de la Reflexividad-Impulsividad evaluada por el Matching Familiar Figures Test-20 (MFFT-20). *Clínica y Salud*.
- Buela-Casal, G., Carretero-Dios, H., De los Santos-Roig, M. y Bermúdez, M.P. (en prensa). Psychometric properties of the Matching Familiar Figures Test's Spanish adaptation. *Journal of Psychoeducational Assessment*.
- Buela-Casal, G., De los Santos-Roig, M. y Carretero-Dios, H. (2000). Reflexividad frente a impulsividad en el rendimiento académico: un estudio longitudinal. *Análisis y Modificación de Conducta, 26* (555-583).
- Buela-Casal, G., De los Santos-Roig, M. y Carretero-Dios, H. (2001b). Propuestas de integración en el estudio de los estilos cognitivos: el modelo de las dos dimensiones. *Revista de Psicología General y Aplicada, 54*, 227-244.
- Buela-Casal, G., De los Santos-Roig, M. y Carretero-Dios, H. (en prensa). La reflexividad-impulsividad como una dimensión continua: validación del siste-

- ma de clasificación de Salkind y Wright (1977). *Revista Latinoamericana de Psicología*.
- Burak, J.A. y Enns, J.T. (Eds.) (1997). Attention, development, and psychopathology. Nueva York: Guilford.
- Butter, E.J. y Jeffcott, R.E. (1982). Prediction of adults's reading performances as a function of auditory and visual cognitive styles and intelligence. *Perceptual and Motor Skills*, 54(2), 359-362.
- Butter, E.J. y Snyder, F.R. (1982). Effect of order of presentation on simultaneous and sequential Matching Familiar Figures Test. *Perceptual and Motor Skills*, 55(3), 1259-1262.
- Cairns, F.D. y Cammock, J. (1978). Development of a more reliable version of the Matching Familiar Figures test. *Developmental Psychology*, 5, 555-560.
- Campbell, S.B. y Douglas, V.I. (1972). Cognitive Styles and responses to the threat of frustration. *Canadian Journal of Behavioural Sciences*, 4, 30-42.
- Cruz, O.M. (1987). *A reflexividade-impulsividade em crianças da idade pré-escolar*. Tesis doctoral no publicada: Universidade do Porto (Portugal).
- Curry, L. (1983). An organization of learning styles theory and constructs. Paper presented at the Annual Meeting of the American Educational Research Association (67th, Montreal, Quebec, April 11-15, 1983).
- Davidson, W.B. y House, W.J. (1978). On the relationship between reflection-impulsivity and field-dependence-independence. *Perceptual and Motor Skills*, 47(1), 306.
- Denney, D.R. (1973). Reflection and impulsivity as determinants of conceptual strategy. *Child Development*, 44, 614-623.
- Dickman, S.J. (1993). Impulsivity and information processing. En W.G. McCown, J.L. Johnson y M.B. Shure (Eds.). *The impulsive client: Theory, research and treatment* (pp. 151-185). Washington, DC: American Psychological Association.
- Douglas, V. y Peters, K. (1979). Toward a clearer definition of the attentional deficit of hyperactive children. En G. Hale y M. Lewis (Eds.). *Attention and the development of cognitive skills*. Nueva York: Plenum.
- Douglas, V.I. (1983). Attention and cognitive problems. En M. Rutter (ed.), *Developmental neuropsychiatry*. Nueva York: Guilford.
- Egeland, B. y Weinberg, R.A. (1976). The Matching Familiar Figures test; a look to its psychometric credibility. *Child Development*, 47, 483-491.
- Egeland, B., Bielke, P. y Kendall, P.C. (1980). Achievement and ajustement correlates of the Matching Familiar Figures Test. *Journal of School Psychology*, 18(4), 361-372.

- Finch, A.J., Saylor, C.F. y Spirito, A. (1982). Impulsive cognitive style and impulsive behavior in emotionally disturbed children. *Journal of Genetic Psychology*, 141(2), 293-294.
- Firestone, P. y Martin, J. (1979). An analysis of the hyperactive syndrome: A comparison of hyperactive, behavior problem, asthmatic, and normal children. *Journal of Abnormal Child Psychology*, 7, 261-273.
- Fuhrman, M.J. y Kendall, P.C. (1986). Cognitive tempo and behavioral adjustments in children. *Cognitive Theory and Research*, 10(1), 45-50.
- García-Pérez, M. y Magaz-Lago, A. (1999). *EMAV 1/2. Escala Magallanes de Atención Visual*. Cruces-Barakaldo: Albor-Cohs.
- García-Sevilla, J. (1997). *Psicología de la atención*. Madrid: Síntesis.
- Genser, B., Häfele, A. y Häfele, M. (1978). Reflectivity and impulsivity: Ability or cognitive style. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 10, 114-123.
- Gilmore, M.C. (1968). *A study of concept formation, reflection-impulsivity and production of alternate uses in nine- and ten-year-old boys and girls from different social classes*. Tesis doctoral no publicada. (citado en Messer, 1976).
- Gjerde, P.F., Block, J. y Block, J.H. (1985). Longitudinal Consistency of Matching Familiar Figures Test Performance from Early Childhood to Preadolescence. *Developmental Review*, 3, 292-302.
- Glow, R.A., Lange, R.V., Glow, P.H. y Barnett, J.A. (1983). Cognitive and self-reported impulsiveness: Comparison of Kagan's MFF and Eysenk's EPQ impulsiveness measures. *Personality and Individual Differences*, 4(2), 179-187.
- Gow, L. y Ward, J. (1980). Effects of modification of conceptual tempo on acquisition of work skills. *Perceptual and Motor Skill*, 50(1), 107-116.
- Gow, L. y Ward, J. (1981). A measure of cognitive style with developmentally disabled adolescents. *Australian Journal of Developmental Disabilities*, 7(4), 181-192.
- Gow, L. y Ward, J. (1982a). The Porteus Maze test in the measurement of reflection-impulsivity. *Perceptual and Motor Skills*, 54(3), 1043-1052.
- Gow, L. y Ward, J. (1982b). Extension of the use of measures of cognitive style to moderately-severely retarded trainees in a field setting. *Perceptual and Motor Skills*, 55(1), 191-194.
- Hagen, J.W. y Wilson, K.P. (1982). Some selected thoughts on attention: a reply to Lane and Pearson. *Merrill-Palmer Quarterly*, 28, 529-532.
- Hall, V. y Russell, W. (1974). Multitrait-multimethod analysis of conceptual tempo. *Journal of Educational Psychology*, 66, 932-939.
- Haskins, R. y McKinney, J. (1976). Relative effects of response tempo and accuracy on problem solving and academic achievement. *Child Development*, 47, 690-696.

- Jackson, A., Farley, G., Zimet, S. y Gottman, S. (1979). Optimizing the WISC-R test performance of low- and high-impulsive emotionally disturbed children. *Journal of Learning Disabilities, 12*, 56-59.
- Jones, A.E. (1997). Reflection-impulsivity and Wholist-Analytic: Two fledglings?... or is R-I a Cuckoo?. *Educational Psychology, 17* (1-2), 65-77.
- Kagan, J. (1965a). Impulsive and reflexive children: Significance of conceptual tempo. En J.D. Krumboltz (Ed.). *Learning and the Educational process*. Chicago: Rand McNally.
- Kagan, J. (1965b). Reflection-impulsivity and reading ability in primary grade children. *Child Development, 36*, 609-628.
- Kagan, J. (1965c). Individual differences in the resolution of response uncertain. *Journal of Personality and Social Psychology, 2*, 154-160.
- Kagan, J. (1966a). Developmental studies in reflection and analysis. En A. Kidd y J. Rivoire (Eds.). *Perceptual Development*. International University Press. New York.
- Kagan, J. (1966b). Reflection-impulsivity: The generality and dynamics of conceptual tempo. *Journal of Abnormal Psychology, 71*, 17-24.
- Kagan, J. (1976). Commentary. En T. Zelniker y W.E. Jeffrey, Reflective and impulsive children: Strategies of information processing underlying differences in problem solving. *Monographs of the Society for Research in Child Development, 41, 5*, 48-52.
- Kagan, J. y Kogan, N. (1970). Individual variation in cognitive processes. En P. Mussen (Ed.). *Carmichael's Manual of Child Psychology* (vol. 1). Nueva York: Wiley.
- Kagan, J. y Messer, S.B. (1975). A reply to "Some misgivings about the Matching Familiar Figures test as a measure of reflection-impulsivity". *Developmental Psychology, 11*, 244-248.
- Kagan, J., Lapidus, D. y Moore, M. (1978). Infant antecedents of later cognitive functioning. *Child Development, 49*, 1005-1023.
- Kagan, J., Moss, H.A. y Siegel, J.E. (1963). Psychological significance of styles of conceptualization. *Monographs of the Society for Research in Child Development, 27* (2), 73-112.
- Kagan, J., Pearson, L. y Welch, L. (1966). Conceptual impulsivity and inductive reasoning. *Child Development, 37*, 583-594.
- Kagan, J., Rosman, B., Day, D., Albert, J. y Phillips, W. (1964). Information processing in the child: Significance of analytic and reflective attitudes. *Psychological Monographs, 78*, (1, número completo, 578).
- Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs: Prentice-Hall.
- Kaufman & Kafuman (1998). *Bateria de Inteligencia General para niños K-ABC*. Madrid: TEA (orig. de 1978).

- Kendall, P. y Wilcox, L. (1979). Self-control in children: Developmental of a rating scale. *Journal of Consulting and Clinical Psychology*, 47, 1020-1029.
- Kennedy, C.B. y Butter, E.J. (1978). Cognitive style in two modalities: Vision and audition. *Journal of Educational Psychology*, 70(2), 193-199.
- Keogh, B.K. y Donlon, G. (1972). Field dependence, impulsivity and learning disabilities. *Journal of Learning Disabilities*, 5, 331-336.
- Kirby, E. (1986). El trastorno por déficit de atención. México: Limusa.
- Kogan, N. (1976). *Cognitive styles in infancy and early childhood*. Hillsdale, NJ: LEA.
- Lane, D.M. y Parson, D.A. (1982). The developmental of selective attention. *Merrill-Palmer Quaterly*, 28, 317-337.
- Larsen, W.W. (1982). The relationship of reflection impulsivity to intelligence and field dependence in older adults. *Journal of Psychology*, 111(1), 31-34.
- Looper, A.B. y Hallahan, D.P. (1980). A comparison of different statistical procedures for determining the relationship between cognitive tempo and reading achievement. *Journal of General Psychology*, 102(1), 89-97.
- López-Soler, C. y García-Sevilla, J. (1997). *Problemas de atención en el niño*. Madrid: Pirámide.
- Manga, D., Fournier, C. y Navarredonda, A.B. (1995). Trastorno por déficit de atención en la infancia. En A. Belloch, B. Sandín y F. Ramos (eds.), *Manual de psicopatología* (vol. 2, pp. 695-718). Madrid: McGraw-Hill.
- Margolis, H. (1982). Conceptual tempo as a moderator variable in predicting first grade achievement test scores. *Journal of School Psychology*, 20(4), 313-322.
- Margolis, H. y Branningan, G.G. (1976). Conceptual impulsivity as a consideration in test interpretation. *Psychology in the Schools*, 13(4), 484-486.
- Margolis, H., Peterson, N. y Leonard, H.S. (1979). Conceptual tempo as a predictor of first-grade reading achievement. *Journal of Reading Behavior*, 10(4), 359-362.
- McCormick, C.B., Miller, G.E. y Pressley, M. (Eds.) (1989). *Cognitive strategy research: From basic research to educational applications*. Nueva York: Springer-Verlag.
- Messer, S.B. (1976). Reflection-Impulsivity: A Review. *Psychological Bulletin*, 83(6), 1026-1052.
- Milich, R. y Kramer, S. (1984). Reflection on impulsivity. An empirical investigation of impulsivity as a construct. *Advances in Learning and Behavioral Disabilities*, 3, 57-94.
- Milich, R. y Kramer, S. (1984). Reflection on impulsivity. An empirical investigation of impulsivity as a construct. *Advances in Learning and Behavioral Disabilities*, 3, 57-94.

- Mitchell, C. y Ault, R.L. (1979). Reflection-impulsivity and the evaluation process. *Child Development*, 50(4), 1043-1049.
- Miyakawa, J. (1981). Some comments on Salkind and Wright's model for reflection-impulsivity. *Perceptual and Motor Skills*, 52, 947-954.
- Moreno, I. (1995). *Hiperactividad: prevención, evaluación y tratamiento*. Madrid: Pirámide.
- Munar, J., Rosselló-Mir, J. y Sánchez-Cabaco (Eds.) (1999). *Atención y percepción*. Madrid: Alianza.
- Navarro, M.J. (1987). *El Estilo Cognitivo Impulsividad-Reflexividad y otras variables del sujeto*. Tesis Doctoral. Facultad de Psicología, Universidad Complutense de Madrid.
- Neimarck, E.D. (1975). Longitudinal development of formal operations thought. *Genetic Psychology Monographs*, 91, 171-225.
- Nisbet, J. y Shucksmith, J. (1987). *Estrategias de aprendizaje*. Madrid: Santillana (original, *Learning strategies*. London: Routledge y Kegan Paul, 1986).
- Oas, P. (1985). The psychological assessment of impulsivity. A review. *Journal of Psychoeducational Assessment*, 3 (2), 141-156.
- Odom, R.D. (1982). Lane and Pearson's inattention to relevant information: A need for the theoretical specification of task information in developmental research. *Merrill-Palmer Quarterly*, 28, 339-345.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Nueva York: Oxford University Press.
- Palacios, J. (1982). Reflexividad-Impulsividad. *Infancia y Aprendizaje*, 17, 29-69.
- Paulsen, K. y Johnson, M. (1980). Impulsivity: A multidimensional concept with developmental aspects. *Journal of Abnormal Child Psychology*, 8, 269-277.
- Peters, R. y Bernsfeld, G. (1983). Reflection-Impulsivity and social reasoning. *Developmental Psychology*, 19, 78-81.
- Pratt, M.W. y Wickens, G. (1983). Checking it out: Cognitive style, context, and problem type in children's monitoring of text comprehension. *Journal of Educational Psychology*, 75(5), 716-726.
- Pressley, M., Borkowski, J.C. y Schneider, W. (1987). Cognitive strategies: Good strategy users coordinate metacognition and knowledge. En R. Vasta y G. Whitehurst (Eds.), *Annals of Child Development*, vol. 4 (89-129). Greenwich, CT: JAI Press.
- Quiroga, M.A. y Forteza, J.A. (1988). La Reflexividad-impulsividad. Estado de la cuestión y análisis de las características psicométricas del test MFF-20. *Investigaciones Psicológicas*, 5, 97-125.
- Rapport, M. D., Chung, K. M., Shore, G., Denney, C. B. e Isaacs, P. (2000). Upgrading the science and technology of assesment and diagnosis: La-

- boratory and clinic-based assesment of children with ADHD. *Journal of Clinical Child Psychology*, 29, 4, 555-568.
- Reales, J.M. y Ballesteros, S. (1997). TDS. Un programa de ordenador para la teoría de detección de señales. Madrid: Editorial Universitas.
- Riding, R.J. (1991). *Cognitive Styles Analysis*. Birmingham (EN): Learning and Training Technology.
- Riding, R.J. y Cheema, I. (1991). Cognitive styles- an overview and integration. *Educational Psychology*, 11, 193-215.
- Rollins, H.A. y Genser, L. (1977). Role of cognitive style i a cognitive task: A case favoring the impulsive approach to problem-solving. *Journal of Educational Psychology*, 69, 281-287.
- Rosselló-Mir, J. (1996). *Psicología de la atención. Introducción al estudio del mecanismo atencional*. Madrid: Pirámide.
- Rosvold, H.E., Mirsky, A.F., Sarason, I., Bransome, E.D., y Beck, L.H. (1956). A continuous performance test of brain damage. *Journal of Consulting Psychology*, 20, 343-350.
- Ruiz-Vargas, J.M. y Botella, J. (1987). Atención. En J.M. Ruiz-Vargas (Dir.), *Esquizofrenia: Un enfoque cognitivo*. Madrid: Alianza.
- Salkind, N.J. (1977). Cognitive tempo in children and their parents. *Perceptual and Motors Skills*, 44(2), 395-400.
- Salkind, N.J. (1977). Cognitive tempo in children and their parents. *Perceptual and Motors Skills*, 44(2), 395-400.
- Salkind, N.J. (1978). Development of norms 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, 237-238.
- Sandberg, S, Rutter, M. y Taylor, E. (1978). Hyperkinetic disorder in psychiatric clinic attenders. *Developmental Medicine and Child Neurology*, 20, 279-299.
- Sanders, A.F. (1983). Toward a model of stress and human performance. *Acta Psychologica*, 53, 61-97.
- Servera, M. & Llabrés, J. (2000). *EMIC: Escala Magallanes de Impulsividad Computarizada (manual y CD-Rom)*. Cruces-Barakaldo: Albor-Cohs.
- Servera, M. (1990). *L'estil cognitiu reflexivitat-impulsivitat i la seva relació amb variables de rendiment i conducta a l'infant*. Tesis de licenciatura no publicada. Universitat de les Illes Balears.
- Servera, M. (1992). *La modificación de la reflexividad-impulsividad y el rendimiento académico en la escuela a partir del enfoque de la instrucción en estrategias cognitivas*. Tesis doctoral. Palma de Mallorca: Universitat de les Illes Balears

- Servera, M. (1997). Evaluación de los estilos cognitivos. En G. Buela-Casal & J. C. Sierra (Eds.), *Manual de evaluación psicológica* (pp. 683-704). Madrid: Siglo XXI.
- Servera, M. (1999). Alteraciones atencionales. En E. Munar, J. Rosselló, & A. Sánchez-Cabaco (Eds.), *Atención y percepción* (pp. 151-174). Madrid: Alianza.
- Servera, M., Bornas, X. & Moreno, I. (2001). Hiperactividad infantil: conceptualización, evaluación y tratamiento. En V.E. Caballo & M.A. Simón (eds), *Manual de psicología clínica y del adolescente* (pp. 401-433). Madrid: Pirámide.
- Servera, M., Bornas, X. y Moreno, I. (2001). Hiperactividad
- Shiffrin, R.M. y Schneider, W. (1977). Controlled and automatic human information processing II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, 88, 127-189.
- Shorr, D.N. y Dale, P. (1984). Reflectivity bias in picture-pointing grammatical comprehension tasks. *Journal of Speech and Hearing Research*, 27(4), 549-556.
- Smith, J.D. y Kemler-Nelson, D.G. (1988). Is the more impulsive children a more holistic processor?. A reconsideration. *Child Development*, 59, 719-727.
- Solís-Cámara R., P. & Solís-Cámara V, P. (1987) Is the matching familiar figures test a measure of cognitive style? A warning for users. *Perceptual and Motor Skills*, 64(1), 59-74
- Sternberg, R.J. (1982). *Handbook of Human Intelligence*. Cambridge University Press.
- Swanson, J.M. (1984). Laboratory learning and attention deficit disorders: use of classical paradigm to assess responses to stimulant medication. En J.M. Swanson, L. Bloomingdale, R. Klorman (Eds.). *Emerging trends in attention deficit disorders*. Nueva York: Spectrum.
- Tanner, W.P. y Swets, J.A. (1954). A decision taking theory of visual detection. *Psychological Review*, 61, 401-409.
- Taylor, E.A. (1986). Déficit de atención. En E.A. Taylor (Ed.), *El niño hiperactivo* (pp. 71-98). Barcelona, Martínez Roca (trad. cast., 1991).
- Taylor, E.A. (1991). Impulsividad, conducta desafiante y problemas de conducta. En E. A. Taylor, *El niño hiperactivo*, Barcelona: Martínez Roca (98-106) (original, "The overactive child", Oxford: Braswell, 1986).
- Taylor, E.A. (1994). Syndromes of attention deficit and hyperactivity. En M. Rutter, E. Taylor y L. Hersov (eds.), *Child and adolescent psychiatry: Modern approaches*. Nueva York: Blackwell Scientific.
- Torner, I., Holstein, R. y Hetherington, E. (1977). Reflection-Impulsivity and self-control in preschool children. *Child Development*, 48, 235-245.

- Tudela, P. (1988). La teoría de detección de señales. En P. Tudela, *Psicología Experimental* (vol. 1, pp. 49-67). Madrid: UNED.
- Vacc, N.A. y Mercurio, S. (1977). Two measures and teacher ratings of impulsivity. *Behavioral Disorders, 2*(2), 114-115.
- Van de Meere, J.J. (1996). The role of attention. En S. Sandberg (Ed.), *Hyperactivity disorders of childhood* (pp. 111-148). Cambridge: Cambridge University Press.
- Vega, J.L. (1984). Desarrollo de los procesos atencionales. En J. Mayor (Ed.), *Actividad humana y procesos cognitivos* (pp. 53-64). Madrid: Alhambra.
- Victor, J.B., Halverson, C.H. y Montague, R.B. (1985). Relationship between reflection-impulsivity and behavioral impulsivity in preeschool. *Developmental Psychology, 21*(1), 141-148.
- Weintraub, S. (1973). Self-control as a correlate of an internalizing externalizing symptom dimension. *Journal of Abnormal Child Psychology, 1*, 292-307.
- Weithorn, C.J., Kagen, E. y Marcus, M. (1984). The relationship of activity level ratings and cognitive impulsivity to task performance and academic achievement. *Journal of Child Psychology and Psychiatry, 25*(4), 587-606.
- Wield, K.H. (1980). Compensatory interventions within the frame work of a diagnostic investigation of intelligence in cognitive-impulsive children (Germ.). *Zeitschrift für Klinische Psychologie Forschung und praxs, 9* (3), 219-231.
- Wright, J.C. (1971). *The Kansas Reflection-Impulsivity scale for Preschoolers. (KRISP)*. St. Louis: CEMREL, Inc.
- Wright, J.C. (1973). *A user's anual for the KRISP_The KIRSP a technical Reports*. St. Louis: CEMREL, Inc.
- Yando, R.M. y Kagan, J. (1968). The effect of teacher tempo on the child. *Child Development, 39*, 27-34.
- Zelniker, T. y Jeffrey, W.E. (1976). Reflective and impulsive children strategies of information processing underlying differences in problem solving. *Monographs of the Society for Research in Child Development, 41* (5 completo).
- Zelniker, T. y Jeffrey, W.E. (1979). Attention and cognitive styles in child. En G.A. Hale y M. Lewis (eds.). *Attention and Cognitive Development*. Nueva York: Plenum.

ISBN 84-369-3502-0



9 788436 935028



MINISTERIO
DE EDUCACIÓN,
CULTURA Y DEPORTE