

DIAGNÓSTICO DE LOS ALUMNOS MÁS DOTADOS

por Carmen JIMÉNEZ FERNÁNDEZ *

Universidad Nacional de Educación a Distancia

1. Introducción

Este estudio se inscribe en una línea de investigación sobre *Diagnóstico y educación de los alumnos más dotados* que se lleva a cabo en el Departamento de Métodos de Investigación y Diagnóstico en Educación de la Facultad de Educación de la UNED y que tiene por finalidad elaborar modelos de diagnóstico y programas educativos específicos para estos alumnos a lo largo de la escolaridad obligatoria, fundamentalmente, extremos sobre los que trabajan varios doctorandos.

La consideración educativa de los chicos más capaces tiene intrínseco interés pedagógico y a él se añade el de la actualidad, adquirida a raíz de la publicación del R. D. 696/1995 y de la Orden y Resolución posteriores que regulan, al fin, la problemática de los alumnos más capaces, tema pendiente en la Reforma y que era una evidente laguna de la Ley de Ordenación General del Sistema Educativo, como hemos señalado en otro lugar (Jiménez Fernández, 1993a).

En las líneas que siguen se ofrece una síntesis de las conceptualizaciones y modelos teóricos dominantes en el campo de la superdotación, la problemática específica del diagnóstico pedagógico de la dotación superior y la de los principales instrumentos en él utilizados. En la segunda parte se ofrecen datos de nuestra

* Con la colaboración de Teresa Aguado Odina, Beatriz Álvarez González, Rosario Jiménez Frías, Enriqueta de Lara Guijarro, Jesús Mesanza López y M^a José Romero Vélez

propia investigación referidos, a los resultados obtenidos por los chicos más dotados en cuatro tipos de pruebas y al comportamiento de dichas pruebas en el diagnóstico de la capacidad. Todo ello con la finalidad de concienciar a profesores, centros y sociedad en general, sobre las peculiaridades de estos alumnos y la obligación de educarlos de acuerdo con su perfil de capacidad.

2. *Conceptualización de la dotación y el talento*

No resulta fácil definir con brevedad y claridad qué se entiende por dotación superior y, sin embargo, es obligado intentarlo dado que su conceptualización manda sobre el diagnóstico y le precede.

Queda lejos la definición dada por Terman en los años veinte que consideraba que un chico era superdotado si pertenecía al uno por ciento superior de la población en inteligencia general. Equivale a una puntuación de al menos 140 puntos en el test de Stanford-Binet o similar. Como estos tests se validaban acudiendo al rendimiento medido a través de pruebas o de las estimaciones del profesor, la definición de dotación quedó muy limitada al potencial académico. A partir de los cincuenta las definiciones amplían el campo y junto al CI y al elevado rendimiento escolar pasan a considerar otras dimensiones.

Witty (1958) es un exponente de ello y considera dotado a todo niño cuyas potencialidades en artes aplicadas, escritura o liderazgo social puedan ser consistentemente reconocidas por sus ejecuciones. Los trabajos de Guilford sobre la estructura de la inteligencia en los años cincuenta inspiraron la propuesta de Taylor (1967) del *Talento múltiple* con seis categorías: académico, pensamiento productivo, comunicación, previsión, toma de decisiones y planificación, que revisaría años más tarde.

Quizá sea la de Marland (1972) la definición más popular al haber sido adoptada por el Departamento de Educación de los Estados Unidos. Dice que los chicos superdotados y con talentos son aquellos identificados por profesionales cualificados y que en razón de sus sobresalientes capacidades son capaces de alta ejecución. Y propone seis categorías: Inteligencia general, aptitud académica específica, pensamiento creativo o productivo, capacidad de liderazgo, habilidades en las artes visuales o representativas y capacidad psicomotora. Abre el campo y advierte de que las capacidades y habilidades de estos chicos es difícil que encuentren satisfacción adecuada en el currículo escolar ordinario.

Renzulli incorpora a la inteligencia y la creatividad, la motivación como componentes esenciales de la dotación superior. La concibe como el resultado de combinar tres variables complejas relativamente independientes: alta inteligencia, creatividad y motivación de logro y propone su modelo de los tres anillos. Critica a Marland por ignorar los factores motivacionales, mezclar categorías de distinta naturaleza, situar demasiado altas las puntuaciones a obtener en los tests de inteligencia y el no dar pautas para desarrollar en la práctica su propia propuesta (Renzulli y Delisle, 1982). Es un modelo bien aceptado si bien se le ha criticado porque podría estar olvidando a los superdotados con mal rendimiento académico dado que la motivación de logro es baja o inexistente en estos alumnos y, sin embargo, es en él un componente esencial (Gagné, 1991), extremo que rechaza su autor (Renzulli, 1994).

La dotación deja de ser algo monolítico y la propia concepción de la inteligencia se diversifica. A su vez la motivación y el contexto social cobran cada vez mayor presencia, Tannenbaum (1985, 1991) considera que la dotación superior se refleja en la vida adulta y que un chico denota potencial para ser un ejecutor críticamente aclamado o productor de ideas ejemplares en campos que engrandezcan la vida moral, física, emocional, social, intelectual o estética de la humanidad. Para llegar a ello deben confluír los factores siguientes: inteligencia general superior, aptitud o aptitudes especiales notables, ordenado soporte de rasgos no intelectivos, ambiente estimulante y facilitador y suerte en momentos cruciales de la vida.

Otra concepción interesante es la de Gardner (1983) y su teoría de las *inteligencias múltiples* (IM). Establece tres principios: 1) La inteligencia no es una dimensión unitaria, una cosa simple sino que la aptitud cognitiva es mejor descrita como un conjunto de capacidades, talentos o habilidades mentales a las que denomina inteligencias. 2) Estas inteligencias son independientes una de otras, y 3) Dichas inteligencias interactúan. Y distingue siete tipos: lingüística, lógica matemática, espacial, musical, físico-kinestésica, interpersonal e intrapersonal, las dos últimas son áreas de comprensión ínter e intrapersonal. Son siete candidatos pero pueden no ser los únicos y cada uno consta de categorías o talentos más específicos (Ramos-Ford y Gardner, 1991).

Otra teoría es la de Gagné (1991) que asocia la dotación superior con el desarrollo natural o no sistemático de las capacidades o aptitudes humanas y con el desarrollo sistemático de los talentos o habilidades que constituyen la pericia en un campo particular de

la actividad humana. Propone el siguiente modelo (Tabla 1) e indica que las diferencias entre aptitudes y talentos no son de naturaleza intrínseca aunque maximizándolas puede decirse que las aptitudes están más ligadas a las características personales, a los procesos o caminos en que se realiza la actividad, a los componentes hereditarios y al potencial de desarrollo; suelen ser descritas mediante la forma adverbial: intelectualmente o físicamente dotado por ejemplo, y se miden de modo aceptable mediante tests estandarizados. Por su parte, los talentos suelen estar más ligados a campos específicos de la actividad humana, al producto de dicha actividad, al ambiente pues en su desarrollo juega un importante papel el entrenamiento sistemático y la actualización del potencial; se les suele describir mediante un complemento directo: está dotado o tiene talento para la música o las matemáticas por ejemplo y son mejor medidos a través de la actividad normal que sirve para definirlos que a través de tests estandarizados

TABLA 1.—*Modelo de superdotación y Talento (Gagné, 1991, 67)*

Una última concepción a destacar es la teoría triárquica de la inteligencia de Stenberg. Consta de tres subteorías: componencial, experiencial y contextual. La primera especifica tres tipos de procesos que subyacen en el procesamiento de la información y que denomina metacomponentes, componentes de ejecución y de adquisición de conocimientos, respectivamente. Dichos componentes han de aplicarse al aprendizaje de tareas nuevas y a interiorizar y automatizar lo aprendido, nos dice desde la subteoría experiencial. Por último hay que adaptarse a un ambiente dado y/o modelarlo. El superdotado posee determinados componentes que no posee el subdotado ni aún la persona normal, aparte de que destaca por su habilidad para combinarlos en caminos no convencionales gracias a su capacidad para mediar con la novedad, captar el conocimiento implícito, su superior capacidad introspectiva, la facilidad para automatizar eficazmente la información nueva y la capacidad de codificación, combinación y comparación selectiva de la información, fundamentalmente. También destacan en la capacidad para gestionar los propios recursos y para compensar sus limitaciones. Dependiendo de los componentes en los que destacan, clasifica a los superdotados en analíticos, sintéticos o creativos y prácticos, si bien sostiene que existe dentro de estos grupos una amplia heterogeneidad (Stenberg, 1991, 1993).

Podrían citarse otros autores pero creemos que los apuntados representan suficientemente el estado de la cuestión.

3. *De la concepción al diagnóstico*

Aunque Terman concibió la superdotación como inteligencia general superior y vio en los tests de cociente intelectual (CI) los instrumentos adecuados para diagnosticarla, en informes posteriores vino a reconocer que, junto a la inteligencia, eran necesarias otras características de la personalidad como el compromiso con la tarea o la motivación de logro para realizar el potencial superior en la vida adulta (Terman, 1970) pero no avanzó en esa dirección.

Por su parte el informe de Marland para el Congreso de Estados Unidos no aborda la cuestión del diagnóstico específico. Bien distinto es el caso de Renzulli, Gardner o Stenberg que han tratado de proyectar en la práctica sus propias teorías a través de programas e instrumentos de evaluación diagnóstica y diferenciación curricular.

Renzulli ha ido perfeccionando su modelo de enriquecimiento. Partidario de una selección no muy restrictiva, propone elegir ini-

cialmente el veinticinco por cien de los alumnos más capaces. Considera que los test de CI y otros medidas de capacidad cognitiva no pueden ser el criterio primordial para identificarlos y propone procedimientos más flexibles que formen parte de la *experiencia educativa ordinaria* y den a los alumnos la oportunidad de participar en una amplia gama de actividades dentro de sus áreas de interés. Por otra parte, dichas actividades permitirán a los profesores observar y registrar el método de trabajo, la satisfacción, compromiso y el producto del alumno y servirán de base para decidir sobre las actividades posteriores de enriquecimiento y profundización curricular.

<p>Su verificación operativizada se refleja en la carpeta del alumno o carpeta del talento que recoge el perfil de aquel en las áreas de capacidad destacadas y sus intereses y sus</p>	<p><i>estilos de aprendizaje</i> se insiste en que se preste la <i>m i s m a atención a las tres dimensiones</i> y se valore al mismo nivel la información obtenida del examen de las realizaciones del estudiante. La Tabla 2 recoge las</p>	<p>delo está siendo aplicado en distintos centros de enseñanza y se le denomina <i>modelo de la puerta giratoria o del triple enriquecimiento</i>. Habilidades Habilidades en ejecución máxima Tests: . Estandarizados . Elaborados por el profesor . Grupos por cursos . Clasificaciones del profesor Evaluación del producto: [1] . Escrito . Oral . Visual</p>	<p>. Musical Intereses Áreas de interés . Artes . Destreza . Literatura . Historia . Matemáticas . Lógica . C. Física . C de la política . Judicial . Atletismo . Recreativas . Márketing . Negocios . Teatro . Danza</p>	<p>. Ejecución musical . Composición musical . Gerenciales . Fotografía . Ordenador . Otros (especificar) Preferencias de estilo Preferencias</p>	<p>de estilo instructivo Recitación y repetición . Tutorías a compañeros . Conferencias . Conferencia/debate . Debates . Estudio independiente guiado . Centro de interés/aprendizaje . Estimulación, roleplaying, dramatización, fantasía guiada</p>
---	---	---	--	--	---

TABLA 2.—Dimensiones de la carpeta de talento (Renzulli, 1994, 186)

- .Juego educativo
- .Informes o pro-yectos para reproducir
- .Informes o pro-yectos de investigación
- .Estudio independiente no guiado
- .Internado
- .Aprendizaje con o sin mentor
- Preferencias en ambiente de aprendizaje

- .Inter/intra-personal:
- .Auto-orientado
- .Orientado al compañero
- Orientado al adulto.
- Combinado
- Física:
- .Sonido
- .Luz
- .Movilidad
- .Momento del día
- .Consumo de comida

Lugar ocupado	
Preferencias en estilo de pensamiento	
Analítico	Se implica fácilmente / rehusa implicarse en la actividad
Sintético/creativo	Muestra confianza / indeciso
Creativo/inventivo	Juguetero / serio
Práctica contextual	Centrado / disperso
Legislativo	Persistente / frustrado por la tarea
Ejecutivo	Capacidad para reflejarse en la tarea / impulsivo
Judicial	Aptitud para trabajar rápido / Aptitud para trabajar lento
Preferencias en estilo de expresión	Comunicación callada
Escrito:	
Oral	Responde a estímulos-indicaciones visuales / auditivos / kinestésicos
Servicio	Es juguetero
Manipulativo	Posee fuerza personal / asuntos para encargar
Debates	Descubre el humor en temas de estudio
Exposición	Usa los materiales de formas sorprendentes
Artístico	Muestra orgullo en el logro de una tarea
Gráfico	Muestra curiosidad por los materiales
Comercial	Le preocupan la corrección de las respuestas
	Se centra en la interacción con el adulto
	Transforma las tareas / materiales

Gadner y su equipo también están aplicando la teoría de las inteligencias múltiples mediante programas dirigidos a los alumnos de la escolaridad obligatoria y de preescolar. La evaluación diagnóstica la integran plenamente dentro de las actividades diarias de la clase y acuden a tareas y estrategias que miden las habilidades de pensamiento y que son evaluadas por referencia criterial. Las denomina *tests auténticos* porque forman parte del trabajo normal del aula. Los resultados en dichas actividades son la base para enriquecer e individualizar el currículo y crear un ambiente que permite a todos los niños descubrir y explorar sus especiales intereses y capacidad. El Proyecto Spectrum para niños de edad preescolar basa el diagnóstico en el examen de siete cam-

pos cognitivos: lenguaje, movimiento, música, números, artes visuales, relaciones sociales y ciencia y en la exploración de estilos de trabajo (Tabla 3).

TABLA 3. —*Medida del estilo de aprendizaje empleado en el Proyecto Spectrum*
(Ramos-Ford y Gardner, 1991, 61)

De acuerdo con su teoría, el diagnóstico de la dotación superior se encamina a observar la capacidad y habilidades del alumno en uno o más campos de los postulados por la teoría de las inteligencias múltiples y trata de identificar la *fuerza relativa* del sujeto en relación con su propio perfil de capacidades cognitivas y la *fuerza absoluta* o situación del alumno en relación con una muestra amplia de pares.

Por último, Stenberg (1993) propone que se asigne la categoría de superdotado apoyándose en cinco criterios: 1) *Criterio de excelencia* según el cual un alumno destaca sobremanera en una dimensión o conjunto de dimensiones al ser comparado con sus iguales. 2) *Criterio de rareza* o que posee un alto nivel en un atributo relativamente raro entre sus iguales. 3) *Criterio de productividad* que establece que la dimensión o las dimensiones en las

que un sujeto es evaluado como superior debe conducir a la productividad real o potencialmente. 4) *Criterio de demostrabilidad* que establece que la dimensión o dimensiones que determinan que un sujeto sea superdotado deben ser demostrables mediante uno o más tests válidos y fiables, y 5) *Criterio de valor*. Es decir que aquello en que destaca un sujeto sea considerado válido para el sujeto y para la sociedad.

Siguiendo con su teoría triárquica ha desarrollado una batería de tests de grupo de niveles múltiples que se puede aplicar desde preescolar hasta adultos y que trata de medirlos distintos tipos de componentes.

Como conclusión podemos destacar: 1) Interés manifiesto por describir qué se entiende por dotación superior y por diferenciar *dimensiones* de dotación con proyección cognitiva, académica y social. 2) Preocupación por *unir teoría y práctica* en el diagnóstico y en la educación de la dotación superior. La dotación debe manifestarse e inferirse en las producciones del alumno en los distintos dominios o ámbitos valorados. 3) Todo diagnóstico tiene que tener clara su *finalidad* y sus objetivos específicos. 4) Situar e integrar el diagnóstico en el contexto natural del centro y aula dándole un *peso similar* a los tests y a las pruebas habituales de clase encaminadas a poner de manifiesto las habilidades cognitivas, intereses y creatividad del alumno. 5) Combinar la evaluación objetiva y subjetiva en una *perspectiva dinámica y complementaria* que mediante la acumulación e interpretación de sucesivas evaluaciones, permita conocer el perfil absoluto y relativo de los alumnos (Jiménez Fernández, 1993b, 1995a).

4. Problemática del diagnóstico

En términos generales los modelos e instrumentos diagnósticos teóricamente han sufrido la misma evolución que la conceptualización pero en la práctica no se refleja satisfactoriamente. Siguiendo a Ritchert (1991) las principales lagunas existentes en este campo son:

a) *Definiciones elitistas y distorsionadas* de la buena dotación. Por ejemplo, se usan definiciones que incluyen sólo ciertos tipos de dotación, normalmente los que menos prevalecen entre los alumnos poco favorecidos social y culturalmente o poco convencionales.

Otras hacen distinciones inadecuadas entre dotación y talento o atribuyen al rendimiento escolar un excesivo poder predictivo de la dotación en la vida adulta.

b) *Confusión acerca de la finalidad del diagnóstico.* Desde nuestra perspectiva debería concebirse como una evaluación de necesidades cuyo propósito es asignar a los estudiantes a programas o, si se prefiere, adaptar el programa al estudiante de modo que desarrolle su potencial en el campo en que destaca. Con frecuencia padres, profesores, directores y orientadores esperan ver en dicha evaluación la confirmación de sus expectativas, prácticas y liderazgo y distorsionan el proceso y/o su finalidad.

c) *Mal uso y abuso de los tests.* Hay centros que emplean como intercambiables los test de cociente intelectual y de rendimiento y que usan en momentos inadecuados determinados instrumentos o procedimientos como acudir a recoger información paterna o a medir la actividad sólo después de que el alumno ha sido seleccionado para un programa de enriquecimiento.

d) *Violación de la equidad educativa.* Las poblaciones social y culturalmente menos favorecidas están escasamente representadas en los programas para biendotados y los instrumentos diagnósticos usados, tests y medidas de rendimiento principalmente, son parcialmente responsables pues penalizan a los alumnos con problemas de rendimiento y dificultades de aprendizaje y a los creativos y críticos.

e) *Distorsión y uso aparente de criterios múltiples.* El intento de acudir a criterios múltiples puede resultar viciado en su raíz ya porque se acude a fuentes no fiables como listas u observaciones sin contraste previo; ya porque se ponderan de tal modo que uno de los criterios es siempre el decisivo; ya porque se combinan medidas de naturaleza distinta, etc. Al utilizar medidas múltiples aumenta la apariencia de buen hacer y suele llevar a una concepción desafortunada de la capacidad superior y a un mantenimiento sostenido del error y la injusticia.

f) *Programas excluyentes y exclusivos.* Ocurre cuando los criterios de admisión a un programa dado son demasiado selectivos de modo que sólo se beneficia al dos o máximo al cinco por ciento de la población escolar, población que por las razones apuntadas tiende a pertenecer a la típica clase media. Y esta situación es un problema para los admitidos y para los excluidos.

Los problemas apuntados existen y aunque sigue primando el diagnóstico mediante tests de inteligencia y de rendimiento escolar,

se aprecia varias tendencias. Una ya apuntada es qué programas y equipos específicos como el liderado por Gardner, han optado por unos instrumentos diagnósticos en el extremo opuesto de los tests psicométricos, en el sentido de que acuden a tests de clase o actividades habituales del programa escolar (Guskin, Peng y Simon, 1992). En segundo lugar el movimiento en torno a los chicos biendotados con problemas de rendimiento escolar (Butler-Por, 1993; Jiménez Fernández, 1996; Jiménez Fernández y Álvarez González, 1997) está promoviendo el desarrollo de estrategias e instrumentos complementarios o alternativos, si bien los tests llamados libres de influencia cultural siguen estando presentes (Clasen, Middleton y Connell, 1994). También se acude con estos alumnos al diagnóstico de la zona de desarrollo potencial apoyándose en las ideas y recursos de Vigostky y Feuerstein (Stanley, 1993). Por último, la mayor parte de los autores recomiendan un enfoque complementario.

Para terminar este apartado señalar que quizá la primera cuestión a plantearse sea el *para qué* del diagnóstico y si se pretende que se incorpore con carácter general a la realidad escolar. Si se pretende iniciar desde preescolar una cierta diferenciación del currículo que alcance también a los chicos de altas capacidades, parece conveniente ensayar estrategias diagnósticas eficaces y que no supongan una complicación y encarecimiento excesivo de la situación escolar. Desde esta perspectiva, puede ser de interés un repaso crítico a los instrumentos diagnósticos más usados en los modelos teóricos.

5. Instrumentos diagnósticos

Se agrupan en dos categorías, tests y juicios de expertos. Estos últimos se apoyan tanto en el análisis de las tareas realizadas por el alumno como en la observación de cómo las realiza y cómo se relaciona con los demás.

5.1. Tests psicométricos

Los tests de inteligencia siguen ocupando un lugar fundamental y en muchos casos es el criterio decisivo para incluir o excluir a un alumno en un programa para superdotados. Se prefieren los denominados libres de influjo cultural tipo WISC, Matrices progresivas de Raven (Matthews, 1993), o K-ABC de Kaufman y Kaufman. Este último parece estar más afectado que otros tests por el efecto techo y los superdotados obtienen en él consistentemente puntuaciones más bajas que en el WISC (Hayden, Furlong y Linne-

meyr, 1988) y en el Stanford-Binet, otros dos tests de uso común con superdotados. El *efecto techo* significa que el alumno no puede obtener la calificación que podría dado que las cuestiones de la prueba se quedan en un nivel bajo de dificultad para su capacidad. En los tests que arrojan varios tipos de puntuaciones como el WISC, estos chicos muestran una amplia dispersión en los subtests y discrepancia entre la inteligencia verbal y la manipulativa (Wilkinson, 1993).

El problema de estos tests es que no están baremados pensando en chicos de alta capacidad sino en los chicos medios y su *fiabilidad* suele ser más baja en la cola superior de la distribución y más elevada en el centro. Para que un test discrimine bien a estos alumnos, debería poseer un suficiente número de elementos por encima de la primera desviación típica para evitar que algunos errores “tontos” hagan bajar la puntuación de modo exagerado, lo que no ocurre en el tramo de la normalidad. Una consecuencia importante del efecto techo es que las puntuaciones muy altas obtenidas en pruebas distintas o a edades diferentes no son comparables. Por otra parte, si el techo de una prueba es demasiado bajo, lo usual es asumir indebidamente que todos los chicos que puntúan por encima de un percentil elevado, digamos del percentil 95, tienen la misma capacidad, supuesto que no se sostiene

En la práctica se tiende a resolver este problema magnificando las pequeñas diferencias de puntuación en el extremo superior de la curva normal o aplicando a estos alumnos un test de dificultad superior, si existe. Cuando se da esta segunda circunstancia, ese pequeño porcentaje de alumnos obtiene en la prueba de nivel superior un amplio rango de puntuaciones, que viene a demostrar que la homogeneidad obtenida en la prueba de su nivel era sólo aparente o forzada por el propio test.

Por otra parte también suelen presentar problemas de *validez*. Aún en la práctica de un test con techo suficientemente alto, en la práctica podrá demostrar su validez sólo si el test o prueba criterio también tiene un techo suficientemente alto. En caso contrario el alumno muy capaz puntuará muy alto en una prueba y menos alto en la otra o criterio y la correlación entre ambas puntuaciones descende. Este problema empalma con la validez de constructo y con el propósito del diagnóstico. Si la finalidad de éste es servir de base a la elaboración de programas educativos adecuados al perfil de capacidad, las pruebas de aptitudes intelectuales y de rendimiento escolar tienen un elevado interés (Serrano, 1992). En esta dirección se mueve la identificación de alumnos con talentos para áreas y programas específicos como las matemáticas (Lupkowski-

Shoplik y Assouline, 1993; Stanley y Benbow, 1986).

Los tests de pensamiento divergente presentan mayores problemas de validez que los de aptitudes y de rendimiento pues difícilmente han sido validados con criterios reales de logros creativos en la población escolar y su validez predictiva resulta problemática.

5.2. *Juicios de expertos*

El recurso a los juicios emitidos por profesores, padres, compañeros y por el propio alumno sigue estando presente en los diferentes modelos sin que esté clara su validez. Aportan una información parcialmente distinta que los tests pues éstos ofrecen resultados comparables entre alumnos de diferentes centros y reflejan su actuación en un momento dado, mientras los juicios son evaluaciones “subjetivas” no comparables que se fijan prioritariamente en el *comportamiento típico o habitual*. Algunos contraponen indebidamente tests con aptitudes y juicios con actitudes.

Suelen emplearse para compensar la deficiente validez de los tests y para evaluar las motivaciones y actitudes de los alumnos pues realmente es el profesor, según sus defensores, el que tiene un conocimiento de lo que es normal, habitual en un nivel y contexto determinado, a diferencia del juicio de los padres que tienen un marco de referencia más incompleto. En cualquier caso, las percepciones y valoraciones del evaluador tienen un fuerte impacto en la evaluación del alumno como ocurre en todos los casos en que se solicita su valoración.

En el caso de los superdotados, la validez de los juicios del maestro ha sufrido avatares distintos. Uno de los trabajos más citados ha sido el de Pegnato y Birch (1959) que sometieron a estudio la *eficacia y eficiencia* de los siguientes métodos o instrumentos: juicio de profesores, pruebas colectivas de CI, tests colectivos de rendimiento, cuadro de honor y creatividad en artes. Definen la eficacia como la razón existente entre el número de superdotados encontrados con un instrumento o método particular y el número “verdadero” de superdotados según el criterio, definido éste por la obtención de una puntuación de 136 mínimo en el test de Stanford-Binet. Y la eficiencia como la razón entre el número de superdotados encontrados usando un método particular y el número total de sujetos investigados con dicho método. Los resultados pusieron de manifiesto un nivel bajo de eficacia y eficiencia de los juicios de los profesores, si bien han sido revisados como se verá en seguida.

A conclusión similar llegaron García Yagüe y su equipo. Escriben que *«Un porcentaje importante de profesores —el 8 por 100— no respondió o lo hizo inadecuadamente y de forma general fueron demasiado rigurosos —sólo citaron del 0,5 al 2 por 100 de los chicos en las cualidades relacionadas con la buena dotación— y variaban exageradamente de unas clases a otras. Más de la mitad de los bien dotados que seleccionamos desde otros criterios les habían pasado desapercibidos, y un tercio de los que citaban, en cambio, eran chicos a todas luces medios o flojos (García Yagüe y col., 1986, 44).*

Renzulli (1979) ha elaborado un cuestionario destinado al profesorado y otros para los alumnos que pretende identificar a los superdotados y que se está adaptando en el Departamento de Métodos de Investigación y Diagnóstico en Educación de la UNED con el fin de analizar su validez relativa. Pues pese a la imprecisión de los juicios del profesorado, los estudiosos no postulan su retirada sino que lo plantean como alternativa complementaria. Y aducen varias razones: no hay tests totalmente válidos; hay actitudes y aptitudes para cuyo diagnóstico no existen instrumentos de evaluación adecuados; pueden dar información sobre una amplia muestra de tareas; combinados con otros instrumentos aumentan cuantitativa y cualitativamente la precisión del diagnóstico y, por último, hay razones de política educativa que aconsejan incluir a profesores y a padres en cualquier modelo diagnóstico.

Por otra parte, la investigación más reciente ha mostrado, de una parte, que las conclusiones del estudio de Pagnato y Birch son inexactas y, por otra, que los profesores aumentan notablemente la precisión y validez de sus juicios si se les da entrenamiento en la dirección deseada. La revisión de Gagné del estudio de Pagnato y Birch pone de manifiesto que los conceptos de eficacia y de eficiencia en él manejados no son independientes. *«Los índices de eficacia y eficiencia son implícitamente tratados por los autores como distintos e independientes... cuando existe una alta correlación negativa entre la eficacia y la eficiencia: de hecho, tal y como ambos conceptos han sido operativizados están directamente relacionados con el número de sujetos empleado en cada método... No se pueden comparar los niveles de eficacia y de eficiencia de cada método puesto que estos dos índices se mueven en direcciones opuestas cuando se varía el punto de corte del criterio (Gagné, 1994, 125).* Conclusión, la calidad y precisión de los juicios del profesorado es mayor de lo que se desprende en este estudio, según Gagné.

Además, la validez de los juicios del docente aumenta con un entrenamiento adecuado. Algunos estudios muestran que los profesores prestan *más significativa* atención a las capacidades y talentos de los alumnos que al género, la clase social o la información étnica si se les expone a descripciones de los talentos y realizaciones de alumnos individuales (Guskin, Peng y Simon, 1992). Esta precisión disminuye en las aulas reales y sobre todo en aquellas en las que la metodología y el contenido es tan homogéneo, que apenas si existe oportunidad de observar y manifestar los talentos específicos.

Otras veces se acude al juicio de los propios compañeros y a la autonominación. Parece que los compañeros son jueces buenos para esta finalidad y sus juicios pueden tener una utilidad máxima en la identificación de superdotados procedentes de culturas y minorías distintas. Gagné defiende esta técnica y señala que la concordancia en el juicio depende primero y fundamentalmente del tipo de aptitud o talento a evaluar, siendo máxima la concordancia entre pares en el caso de las aptitudes intelectual y física (Gagné, Bégin y Talbot, 1993). Clasen, Middleton y Conell (1994) concluyen que los juicios de los estudiantes son altamente precisos en el diagnóstico de la resolución de problemas y de habilidades artísticas.

PREESCOLAR		2.º EGB		6.º EGB		TOTAL	
Masc.	Fem.	Masc.	Fem.	Masc.	Fem.	Masc.	Fem.
82	84	90	111	34	43	208	236
166		201		77		444	

TABLA 4.—Distribución de la muestra total antes de seleccionar a los más dotados

PREESCOLAR				2.º EGB				6.º EGB				TOTAL			
Total 25% 17% 10%				Total 25% 17% 10%				Total 25% 17% 10%				Total 25% 17% 10%			
166	47	30	16	221	50	36	18	71	21	13	7	444	118	79	41
166				201				77				444			

La autonominación tiene sentido a partir de los doce años aproximadamente y no debe plantearse en abstracto sino en relación con determinadas propuestas pragmáticas que permita a los interesa-

dos evaluar si podrían seguirlo y si estarían dispuestos a hacerlo. Renzulli la emplea en su modelo y, como otros autores, le da un valor complementario y no sustitutivo de otras técnicas. No parece se haya investigado sobre su validez.

Sin entrar en nuevas consideraciones sobre los instrumentos diagnósticos remitimos al lector interesado al trabajo de Callahan, Lundberg y Hunsaker (1993) que recoge, en una escala, posibles criterios a considerar en la evaluación o selección de dichos instrumentos de acuerdo con sus objetivos específicos y la población a que se destinan.

6. Estudio experimental

A continuación se presentan resultados de una investigación uno de cuyos objetivos el diagnóstico de los alumnos más dotados

la elaboración de instrumentos según	partiendo de un modelo sólido y operativo, que combine distintos criterios de calidad, como paso previo a cuadrarse en los	diagnóstico no penaliza a ningún tipo específico de alumnos o no lo penaliza más que a otro grupo. c) Se busca la eficacia general pero con el mínimo coste. Ya se ha indicado que
programas de enriquecimiento ordinarios para atender a todas las diferencias individuales con significativamente: a) Estado para la diagnóstico de la educación. b) Dicho dotación debe en-	centros escolares	concebimos la dotación como un constructo complejo, multidimensional y dinámico que debe ser detectada, estimulada y valorada por la educación sistemática.

concebimos la dotación como un constructo complejo, multidimensional y dinámico que debe ser detectada, estimulada y valorada por la educación sistemática.

Se ha trabajado con alumnos de preescolar y de segundo y sexto curso de Educación General Básica, pertenecientes a dos centros públicos y dos centros privados de Madrid capital y a un centro público de Zaragoza capital. En total cinco centros que contabilizan 24 aulas, de las cuales 8 pertenecen al segundo curso de preescolar; 7 a segundo de EGB y 9 a sexto de EGB. Cada aula cuenta aproximadamente con una media de 26 alumnos, excepto el preescolar de uno de los centros en el que había más de 30 alumnos por aula.

6.1. Muestra

Consta inicialmente de 444 alumnos de los cuales algo menos de la mitad son varones. Su distribución por sexo y curso se recoge en la Tabla 4. La Tabla 5 recoge la distribución por curso de la muestra general y de las tres submuestras de los más dotados intelectualmente, según que el punto de corte se establezca en el 25%, 17% y 10% superior de la distribución, respectivamente. En efecto, a fin de observar mejor el comportamiento de los datos, siempre difíciles de obtener, y puesto que no existe un criterio unívoco que señale cuándo se considera especialmente dotado a un alumno, hemos optado por analizar los datos con criterios más y menos restrictivos, a sabiendas de que estas submuestras no son independientes. Es decir, la muestra general contiene a las tres restantes; la del 25% incluye a las del 17% y 10% y la penúltima contiene a la última.

TABLA 5.—Distribución de la muestra estableciendo tres cortes en el test criterio (MSCA o WISC)

6.2. Procedimiento e instrumentos

Los instrumentos de diagnóstico utilizados en cada uno de los cursos se recogen en la Tabla 6. El procedimiento tuvo dos varian-

CURSO MUESTRA	PREESCOLAR (MSCA)		SEGUNDO (WISC)		SEXTO (WISC)	
Total	\bar{X}	55,23*	108,41**	103,97	119,56	
	σ	9,05	14,50	13,93	11,93	
	n	166	166	201	77	
25%	\bar{X}	65,83	125,44	121,78	133,71	
	σ	5	8,04	9,05	7,42	
	n	47	47	50	21	
17%	\bar{X}	67,91	128,78	125,11	135,29	
	σ	4,8	7,7	8,58	7,41	
	n	32	32	36	13	
10%	\bar{X}	71,56	134,55	131,83	143,29	
	σ	4,31	7,20	7,20	2,36	

tes: a) En primer lugar, aplicación de la prueba colectiva de inteligencia TEA-2 a los alumnos de sexto curso, prueba que sirvió para efectuar un *cribaje* previo. Se seleccionó al 30% superior de los alumnos en dicha prueba. A esta muestra se sumaron los alumnos nombrados por el profesor (*selección* del profesorado), coincidieran o no en el diagnóstico ambos medios, y al conjunto resultante se les pasó el resto de las pruebas. b) La segunda forma consistió en

<p>aplicar el test individual de inteligencia y seleccionar a los alumnos del 25% superior. A este grupo</p>	<p>del 25% se sumaron, en porcentaje similar como máximo, los alumnos seleccionados independientemente por el profesor, hubieran sido o no seleccionados por el test, y al conjunto resultante se le aplicó el resto de las pruebas (Tabla 6).</p>	<ul style="list-style-type: none"> . Autoconcepto . Percepción del profesor . Rendimiento (intermediato y diferido) . Selección del profesor . WISC . Autoconcepto . Percepción del profesor . Rendimiento (intermediato y diferido) 	<p>profesor WISC</p>
	<p>CURSO PREESCOLAR 2.º EGB 6.º EGB</p>	<ul style="list-style-type: none"> . WISC . Autoconcepto . Percepción del profesor . Rendimiento (intermediato y diferido) 	<p>Los tests se aplicaron en el primer trimestre del curso 1992-93 y no incidiremos en ellos</p>
	<p>INSTRUMENTOS</p> <ul style="list-style-type: none"> . Selección del profesor. . MSCA 	<ul style="list-style-type: none"> . TEA-2 . Selección del 	
	<p>6.3. Resultados</p>		<p>El cuestionario de autoconcepto consta de veintidós preguntas dirigidas a los alumnos y su finalidad es recabar información sobre la imagen que re-</p>

por ser conocidos. A mediados de dicho trimestre se pidió a los profesores que *seleccionaran* a los alumnos más sobresalientes e *interesantes* en características valoradas por la escuela y la sociedad sin limitarse a un número específico y sin informarles de la concreta de su elección. A comienzos del tercer trimestre se pasaron los cuestionarios

de autoconcepto y de percepción del profesor. A finales del curso se recogieron las calificaciones escolares o *rendimiento* intermediario. De una parte de la muestra se recogió el rendimiento obtenido un año escolar más tarde o *rendimiento* diferido.

El cuestionario de autoconcepto consta de veintidós preguntas dirigidas a los alumnos y su finalidad es recabar información sobre la imagen que re-

Se presentan resultados referidos por una parte, a las características de las cuatro muestras en las variables estudiadas y, por otra, referidos a la validez de los instrumentos diagnósticos utilizados cuando el objetivo es la identificación de la capacidad superior.

6.3.1. Características de los diferentes grupos

Como era de esperar, hay importantes diferencias entre la muestra general y los subgrupos superiores, favorables a estos últimos en la totalidad de las características evaluadas. Pueden apreciarse en las tablas 7, 8 y 9 las referidas al test de inteligencia en sus variantes de cociente intelectual total, verbal y manipulativo, respectivamente.

En esta tabla y en las dos siguientes, la primera columna marcada con un asterisco (*), recoge las puntuaciones según los baremos del MSCA, y la segunda marcada con dos asteriscos (**), las puntuaciones transformadas a una escala similar a la del WISC, ello en las cuatro muestras.

TABLA 7.—*Resultados obtenidos en el CI total*

	CURSO MUESTRA	
	Total	
	25%	
	17%	
	10%	
	PREESCOLAR (MSCA)	
\bar{X}	51,01	101,62
σ	8,68	13,96
n	166	166
\bar{X}	61,07	117,70
σ	5,60	9,01
n	47	47
\bar{X}	63,57	121,80
σ	5,48	8,80
n	32	32

X	66,89	127,14
σ	4,63	7,43
n	16	16

SEGUNDO (WISC)

99,92
14,71
201

118,40
9,46
50

122,59
9,24
36

129,56
7,32
18
SEXTO (WISC)

112,51
14,11
77

130,48
9,36
21

135,62
8,34
13

142,14
5,40
7

TABLA 8.—*Resultados obtenidos en el CI Verbal*

CURSO
MUESTRA

Total

25%

17%

10%
PREESCOLAR (MSCA)

\bar{X}	60,84	117,42
σ	10,47	16,83
n	166	166

\bar{X}	74,14	138,80
σ	3,63	5,84
n	47	47

\bar{X}	76,10	141,95
σ	2,02	3,25
n	32	32

\bar{X}	77,75	144,60
σ	0,45	0,72
n	16	16

SEGUNDO (WISC)

105,59
12,91
201

120,72
7,90
50

124,89
7,03
36

130,35
4,65
18

SEXTO (WISC)

120,61 11,39	77	5,99 21	138,15 5,06 13	142,10
3,70 7	135,05		TABLA 9.—	Resultados

obtenidos en el CI manipulativo, entre el CI del grupo total y el de los subgrupos superiores establecidos, se producen diferencias en torno a los 20-30 puntos, tendiendo a ser éstas mayores en el CI manipulativo. Según el WISC, el corte que separa al grupo del 25% superior se sitúa en torno a los 120 puntos en el caso del segundo curso y en torno a los 130 en el caso de sexto curso. Si nos ceñimos al MSCA aplicado en preescolar, observamos que existe una diferencia de 10, 13 y 16 puntos, respectivamente, entre la medida del grupo total y las medidas de los tres subgrupos superiores, respectivamente, si se consulta la puntuación original del baremo. Si se observa la puntuación ya transformada del MSCA, las diferencias son similares a las que arroja el WISC.

Como en investigaciones anteriores, en la nuestra *los chicos más dotados destacan no sólo en inteligencia general sino también en rendimiento escolar inmediato y diferido y en autoconcepto*, mejorando en las dos últimas características conforme aumenta el CI y ello en los tres cursos estudiados. No obstante, en la muestra de segundo curso desciende el rendimiento académico diferido o en tercer curso de algunos de los alumnos seleccionados en el corte del 25% superior. Asimismo *tienen significativamente mayores probabilidades de ser designados por sus profesores como alumnos valiosos o con capacidades*.

Con independencia de la prueba de inteligencia individual utilizada, los cuatro grupos *muestran en los tres cursos estudiados mayor inteligencia manipulativa que verbal*. Sin embargo, *conforme se es más estricto en la selección de la muestra y se asciende en el curso estudiado, la superioridad del cociente manipulativo frente al verbal se mantiene pero desciende su cuantía y llega a desaparecer en el subgrupo del 10% superior de sexto curso*. Se aprecia

esta tendencia al comparar los resultados de cada una de las cuatro muestras de sexto curso. Entre el WISC verbal y el manipulativo existe una diferencia de 8 puntos en la muestra general; de casi 5 puntos en la del 25%; de unos 3 puntos en la del 17% y no existe diferencia en el caso del 10% superior (Tablas 8 y 9).

Si consideramos el *curso*, las *diferencias* obtenidas en la *calificación global* del test, son *altamente significativas* siendo los alumnos de sexto los que obtienen un promedio significativamente superior al obtenido por los de segundo y los de preescolar; los alumnos de segundo curso obtienen la media más baja. En el caso de los subgrupos del 25%, 17% y 10% superior, el análisis de varianza arroja diferencias significativas a favor del sexto curso, siendo el segundo curso el que presenta la media más baja aunque

~~no es significativa respecto de preescolar y segundo curso.~~ Si nos ceñimos a la *inteligencia verbal* y seguimos considerando el curso, *existen diferencias significativas* entre la media obtenida en cada uno de los cuatro subgrupos considerados, siempre a favor de sexto curso. Finalmente, *la inteligencia manipulativa* también arroja *diferencias significativas* a favor de preescolar y sexto frente a segundo curso en las cuatro submuestras consideradas, siendo en todas ellas la media más alta la obtenida en preescolar, como se ha puntuado.

Los alumnos situados en el 25% superior según el test individual, tienen significativamente más probabilidades que los alumnos restantes de ser señalados por sus profesores como alumnos destacados, tanto si se analizan los datos globalmente como si se tiene en cuenta el curso que estudian. Ello viene a indicar que los profesores detectan esa superioridad, aunque de modo imperfecto, pues esta situación se repite cuando la observación del profesor viene guiada por las descripciones de un cuestionario como el de Renzulli.

En el cuestionario de percepción del profesor, los *alumnos de capacidad superior* tienen *significativamente mayores probabilidades* que sus compañeros del 75% inferior *de ser percibidos favorablemente por sus profesores* y *destacan* en dos de las características medidas por dicho instrumento: *planificación y liderazgo*, aumentando sensiblemente la primera desde preescolar a sexto y descendiendo ligeramente la segunda.

Para la muestra total, el autoconcepto mejora de preescolar a sexto en los tres cursos considerados, existiendo diferencias significativas a favor de la muestra de sexto curso. En los tres subgrupos superiores se aprecia la tendencia a un autoconcepto superior al de la muestra general y a una discreta mejora del mismo de preescolar a segundo, descendiendo ligeramente en sexto, siendo significativas las diferencias a favor de segundo curso (Tablas 10 y 11).

El subgrupo del 25% también tiene significativamente más probabilidades de obtener buen rendimiento que el 75% restante, y ello a lo largo de los tres cursos. Por otra parte, *el rendimiento aumenta a medida que el subgrupo es más superior o restrictivo.* Esta tendencia no puede hacer ignorar que existen alumnos con alta capacidad y bajo rendimiento y que este fenómeno parece afianzarse con el paso del tiempo pues el rendimiento académico diferido o en tercer curso desciende en la submuestra de segundo curso del 25% superior. Este extremo había sido apuntado en el estudio de García Yagüe y colaboradores (1986).

El control de la variable *sexo o género* pone de manifiesto que *no existen diferencias significativas entre los géneros en ninguno de los tres CI analizados: total, verbal o manipulativo considerando cada uno de los cuatro grupos establecidos. Tampoco aparecen diferencias ligadas al género en autoconcepto ni en la selección libre del profesor. Sí aparecen diferencias significativas en la escala de planificación y en la de liderazgo cuando el profesor selecciona guiado por el cuestionario de Renzulli, en el primer caso a favor de los chicos y en el segundo a favor de las chicas. También aparecen diferencias significativas en rendimiento académico a favor del sexo femenino.*

Sin entrar en más pormenorizaciones, nuestros datos soportan el retrato robot de los alumnos más dotados como grupo que los describen como alumnos con alta inteligencia, buen rendimiento, autoestima y apreciados por sus profesores, si bien existen subgrupos con problemas de rendimiento escolar y baja autoestima (Jiménez Fernández, 1996; Jiménez Fernández y Álvarez González, 1997), fenómeno que estudiamos en otros trabajos.

6.3.2. *Valoración de los instrumentos empleados*

Bajo este epígrafe se hacen sugerencias sobre la utilidad de los instrumentos diagnósticos por nosotros empleados para identificar a los chicos más capaces.

El test MSCA tiene un techo muy bajo para dicho objetivo po-

niéndose *especialmente* de manifiesto en la *escala manipulativa*, debido a la tendencia a calificar más alto en ella que en la verbal en la edad de *preescolar*. En efecto, en la muestra del 10 % superior de preescolar, la desviación típica es de 0.45, habiendo sólo una diferencia de un punto entre los valores mínimo y máximo obtenido por los 16 alumnos de la muestra, siendo el máximo, 78, el techo de la prueba, techo que es alcanzado por 12 de los 16 alumnos. En la submuestra del 17%, la desviación típica es de 2.02, extremos que pueden apreciarse en la primera columna de las tablas 7, 8 y 9. En el caso del WISC, la desviación típica más pequeña en los ocho grupos analizados es de 3.7.

Por otra parte, se toma como *criterio la prueba individual de inteligencia* para decidir sobre el grado de eficacia y eficiencia de los instrumentos empleados en cada uno de los cursos. La elección de esta prueba tiene como finalidad tener un referente con el que establecer comparaciones. Pero en el caso de que una técnica o instrumento resulte poco eficaz, no debe interpretarse como desaconsejable o poco útil en términos absolutos. El estudio profundo de esta falta de eficacia puede arrojar luz sobre las peculiaridades que penalizan la elección de determinados alumnos según se utilicen unos instrumentos de detección u otros.

En los tres cursos *nos ceñimos al grupo del 17% superior en el criterio* y tomando a continuación a cada una de las pruebas por separado, se ha calculado la eficacia, o cociente entre el número de chicos seleccionados por un instrumento determinado y el número de superdotados seleccionados por el criterio, y la eficiencia, o razón entre el número de biendotados encontrados usando un método particular y el número total de sujetos investigados con dicho método. Como se ha dicho no son índices independientes.

INSTRUMENT.	MUESTRA		MUESTRAS TOTALES
	ALUMNOS	SELECCIONA.	
Prsc. 2° 6°			EFICACIA
Prsc. 2° 6°			(%)
Prsc 2° 6°			EFICIENCIA
Prsc. 2° 6°			(%)
			PRUEBA INTELIGENC.
			CRITERIO
166 201 77			
32 36 13			

		SELECCIÓN											
PROFESOR	166	201	77	13	20	11	41	55	85	25	27	—	
RENDIMIENTO.													
109	94	132	11	15	7	34	42	54	31	45	41		
		AUTO											
CONCEPTO	52	88	69	5	6	4	16	17	31	31	19	45	
		PERCEPCIÓN											
PROFESOR	24	32	42	2	5	4	6	14	31	25	44	74	

TABLA 10.— *Eficacia y eficiencia de los instrumentos empleados en el grupo del 17% superior en el test individual*

Los medios que han demostrado ser más eficaces o que han detectado el mayor número de alumnos previamente seleccionados por la prueba, son, en primer lugar, la selección del profesor, y en segundo lugar, las calificaciones, siendo la selección el instrumento más eficaz en los tres cursos, destacando de manera especial en el caso de sexto curso. El juicio o selección previa del profesor no parece desviarse demasiado de las notas que otorga a fin de curso. La eficiencia de ambos medios tiende a invertirse aunque en el conjunto de los datos mantienen unos niveles interesantes.

Por otra parte, los dos cuestionarios elaborados por Renzulli ofrecen unos resultados poco satisfactorios en los dos índices calculados, lo que en parte se explica por el pequeño tamaño de la muestra. En cualquier caso, parece conveniente revisarlos aunque sea mínimamente, pues contienen expresiones y algunos elementos tal vez demasiado centrados en la cultura escolar norteamericana. El conjunto de los instrumentos han obtenido un mejor funcionamiento en sexto curso que en los dos cursos restantes.

6.3.3. Concordancia entre dos o más medios

En este caso consideramos solamente aquella parte de la muestra que es seleccionada por un determinado número de instrumentos independientes al mismo tiempo siendo *excluyentes* las categorías consideradas. Así, los alumnos seleccionados por cuatro instrumentos al mismo tiempo, se excluyen de los seleccionados por tres, éstos de los de dos y así sucesivamente. Nos ceñimos al grupo del 17% superior. Por otra parte, al tener que coincidir en un número de pruebas, cambia el tamaño de las muestras.

Si acudimos a la Tabla 11 se aprecia: a) Solamente en sexto curso aparecen alumnos seleccionados por los cinco instrumentos al mismo tiempo, si bien en porcentaje muy pequeño: 4.25 por ciento. b) Sobre el 2% de los alumnos de preescolar y segundo curso es seleccionado al mismo tiempo por cuatro de los instrumentos diagnósticos, elevándose al 18.5 % en el caso de sexto curso. c) Entre el 10 y el 15 por ciento de los alumnos son diagnosticados al mismo tiempo por tres pruebas, oscilando entre el 30 y el 35 por ciento en el caso de dos.