

Sumario

Editorial	3
ESTUDIO MONOGRÁFICO	
Algunos retos metodológicos <i>Iñaki Dendaluze Seguro</i>	7
TRABAJOS DE INVESTIGACIÓN	
Elaboración de un sistema de categorías para el análisis sociométrico: su aplicación en aulas multiculturales <i>Flor Cabrera, M^º. Ángeles Marín, Julia Victoria Espín y Mercedes Rodríguez</i>	27
La búsqueda de factores protectores del fracaso escolar en niños en situación de riesgo mediante un estudio de casos <i>Judit Fullana Noell</i>	47
La escala Keymath-R. Estudio de una prueba diagnóstica con referencia curricular <i>Ricard Marí Mollà</i>	71
Un examen de la validez externa del Stat (Sternberg Triarchic Abilities Test) <i>Ángela Rojo, M^º. Dolores Prieto y Juan L. Castejón</i>	85
Determinantes del progreso en los estudios de doctorado. Diferencias entre disciplinas científico-tecnológicas versus ciencias sociales y humanidades <i>Esperança Villar</i>	101
Concordancia y errores (tipo I y II) en el diagnóstico logopédico primario <i>Antonio Fernández Cano y Mariano Machuca Aceituno</i>	123
Propuesta metodológica para la evaluación del gasto educativo de los hogares, en educación reglada no universitaria. Aplicación a la Comunidad Autónoma de la Región de Murcia <i>Juan Gómez, Matilde Lafuente y Antonio Losa</i>	137
Análisis de estructuras de covarianza en el estudio de la Competencia Docente del Profesor Universitario <i>José Manuel García Ramos</i>	155
TRABAJOS METODOLÓGICOS	
Comparación de métodos estadísticos para la evaluación del funcionamiento diferencial de los ítems <i>Javier Gil Flores, Eduardo García y Gregorio Rodríguez</i>	187
FICHAS RESUMEN DE INVESTIGACIONES	207



Revista de Investigación Educativa

Volumen 16, número 1, 1998

Volumen 16, número 1, 1998

Revista de Investigación Educativa

Asociación Interuniversitaria de Investigación Pedagógica (AIDIPE)



ISSN: 0212-4068
Depósito Legal: MU-724-1996

Revista de Investigación Educativa

Volumen 16, número 1, 1998

Asociación Interuniversitaria de Investigación Pedagógica (AIDIPE)

Director:

Juan Mateo Andrés

Directora ejecutiva:

Julia Victoria Espín

CONSEJO ASESOR:

Margarita Bartolomé

Nuria Borrell

Leonor Buendía

Iñaki Dendaluce

José Cajide

Narciso García

Fuensanta Hernández Pina

Jesús Jornet

Mario de Miguel

Arturo de la Orden

Ramón Pérez Juste

Antonio Rodríguez Diéguez

Francisco J. Tejedor

CONSEJO DE REDACCIÓN:

Manuel Álvarez

Rafael Bisquerra

Flor Cabrera

Inmaculada Dorio

Julia Victoria Espín

Pilar Figuera

M^a Ángeles Marín

M^a Luisa Rodríguez

Mercedes Rodríguez

M^a Paz Sandín

Trinidad Donoso

SUSCRIPCIÓN E INTERCAMBIO CIENTÍFICO:

Dpto. de Métodos de Investigación y Diagnóstico en Educación

Facultad de Pedagogía

Paseo del Valle de Hebrón, 171, 2^a planta

08035 BARCELONA (Spain)

DISTRIBUCIÓN:

Dpto. de Métodos de Investigación y Diagnóstico en Educación

Facultad de Educación

Campus de Espinardo

Universidad de Murcia 30100

Tel. (968) 36 40 67

ENTIDADES COLABORADORAS: Ver pág. 2

DISEÑO, MAQUETACIÓN E IMPRESIÓN:

Compobell, S.L. MURCIA

ISSN: 0212-4068

Depósito Legal: MU-724-1996

Revista de Investigación Educativa

Volumen 16, número 1, 1998

Editorial	3
ESTUDIO MONOGRÁFICO	
Algunos retos metodológicos	7
<i>Iñaki Dendaluze Segurola</i>	
TRABAJOS DE INVESTIGACIÓN	
Elaboración de un sistema de categorías para el análisis sociométrico: su aplicación en aulas multiculturales	27
<i>Flor Cabrera, M^a. Ángeles Marín, Julia Victoria Espín y Mercedes Rodríguez</i>	
La búsqueda de factores protectores del fracaso escolar en niños en situación de riesgo mediante un estudio de casos	47
<i>Judit Fullana Noell</i>	
La escala Keymath-R. Estudio de una prueba diagnóstica con referencia curricular	71
<i>Ricard Marí Mollà</i>	
Un examen de la validez externa del Stat (Sternberg Triarchic Abilities Test) ...	85
<i>Ángela Rojo, M^a. Dolores Prieto y Juan L. Castejón</i>	
Determinantes del progreso en los estudios de doctorado. Diferencias entre disciplinas científico-tecnológicas versus ciencias sociales y humanidades	101
<i>Esperança Villar</i>	
Concordancia y errores (tipo I y II) en el diagnóstico logopédico primario	123
<i>Antonio Fernández Cano y Mariano Machuca Aceituno</i>	
Propuesta metodológica para la evaluación del gasto educativo de los hogares, en educación reglada no universitaria. Aplicación a la Comunidad Autónoma de la Región de Murcia	137
<i>Juan Gómez, Matilde Lafuente y Antonio Losa</i>	
Análisis de estructuras de covarianza en el estudio de la Competencia Docente del Profesor Universitario	155
<i>José Manuel García Ramos</i>	
TRABAJOS METODOLÓGICOS	
Comparación de métodos estadísticos para la evaluación del funcionamiento diferencial de los ítems	187
<i>Javier Gil Flores, Eduardo García y Gregorio Rodríguez</i>	
FICHAS RESUMEN DE INVESTIGACIONES	207

E D I T O R I A L

Normalmente no tenemos una sensación concreta respecto del paso del tiempo sino es porque algún referente nos facilita la tarea de situarnos en él. Posiblemente sea, porque no es el tiempo el que pasa sino nosotros, o mejor dicho aquello que más nos sustantiva como seres, es decir: nuestros pensamientos, nuestros planteamientos, nuestras ideas, nuestras actividades, nuestros recuerdos...

Un referente imposible de ignorar es el que inexorablemente se producirá y se viene anunciando desde todos los ámbitos sociales, el cambio de siglo y de milenio. Parece que un hecho puntual de tanta relevancia, se debe convertir en punto de inflexión obligado para replanteamientos de colectivos que como el nuestro ha crecido y se ha desarrollado alcanzando un importante grado de madurez a lo largo de algo que pronto denominarán como «otros tiempos».

Se intuye que todos los ámbitos del saber precisarán practicar un obligado ejercicio de adaptación, para atender a las nuevas necesidades y de reconceptualización para integrar los nuevos conocimientos y metodologías. No afrontar los nuevos retos conducirá irremediabilmente a perder la consideración de personas de nuestro tiempo.

Con el siglo también acaban o al menos se cuestionan muchos de los mitos forjados desde la categoría histórica denominada modernidad y que podríamos resumir en las ideas de progreso y razón forjadas por los comerciantes y humanistas del Renacimiento y por los filósofos y políticos de la ilustración. Ideas que en lo social y en lo científico prometían progreso ilimitado, racionalidad y emancipación.

La crisis de estos principios se produce, tal como señala Lyotard cuando la realidad de final de siglo nos muestra más que el paraíso prometido, la eclosión de nuevas miserias y el reforzamiento e intensificación de los mecanismos de control.

Crece el sentimiento de que posiblemente la historia no es una línea recta y ascendente, unidireccional e irreversible, que conduce necesariamente a un punto absoluto, sino una suerte de espiral recurrente y rítmica, pluridireccional y reversible que puede devolvernos a cualquier situación anterior.

Estas nuevas percepciones de la realidad, no deben conducirnos al pesimismo o al «todo vale» del pensamiento postmoderno, sino hacernos conscientes de la necesidad de imprimir a nuestras acciones intención, voluntad, continuidad, sentido y significado y a no confiar irresponsablemente en aquello de que «el tiempo todo lo arregla».

Nuestro colectivo no puede ignorar la situación y debemos vertebrar nuevas formas organizativas, arbitrar nuevas vías de expresión, promover nuevos esfuerzos

del pensamiento, debemos, en definitiva, generar un ecosistema cultural que garantice la tensión necesaria para producir los cambios necesarios y en la dirección más adecuada.

Sólo desde ese esfuerzo colectivo es posible alcanzar lo que Throsby denomina «sostenibilidad cultural» entendida como el grado de masa crítica cultural preciso para mantener la capacidad de transformación y creación de un ámbito del conocimiento.

Alguien nos comentó en el último Congreso de Sevilla, ¿Es AIDIPE una agencia destinada a gestionar congresos y revistas o es algo más? La respuesta situada en el futuro más inmediato se nos antoja muy simple: o conseguimos ser algo más o simplemente no seremos.

ESTUDIO MONOGRÁFICO

ALGUNOS RETOS METODOLÓGICOS

*Iñaki Dendaluze Seguro*la*

Sin ser milenaristas la proximidad del año 2000 representa una oportunidad para hacer algunos balances y plantearse algunas cosas que deseamos cambiar o introducir en nuestros ámbitos de trabajo profesional. En este artículo se va a dar prioridad a los retos sobre los balances. Naturalmente la lista y discusión de los retos metodológicos de la investigación en Ciencias Sociales varía según quién la exponga y siempre es reflejo de un momento concreto. A veces se presentan los retos en forma de planteamientos de temas a discutir y de posicionamientos sobre los mismos. En todo caso esta propuesta de retos quiere servir para que siga la tarea colectiva de diálogo y contrastes varios.

Para contextualizar algunos de esos retos voy a comenzar con un ejemplo y hablando en primera persona. El ejemplo es el de una investigación que estamos realizando, entre otras de otro contenido y tipo, tres personas (Francisca Arbe, Yosu Artieda e Iñaki Dendaluze) con distintas titulaciones y los tres con experiencia en el mundo real de la educación. Lo de reflexionar con datos sobre nuestro quehacer profesional nos convence a los tres como una forma de contribuir a mejorarlo.

Uno de nosotros no ha querido nunca trabajar en la universidad y tiene acumulada una gran riqueza de datos sobre alumnos de primaria y secundaria, desde que los alumnos tienen 10 años hasta los resultados de sus pruebas de selectividad, desde 1981 hasta el presente.

En esta investigación decidimos centrarnos en los datos de un solo centro y en datos más bien cuantitativos, procedentes en gran parte de pruebas colectivas. Hasta ahora en esta investigación no nos hemos atado a un financiador, ni tenemos plazos fijos para la entrega de resultados y en parte por eso hemos sido muy libres para tra-

* Doctor en Psicología (Medición y Evaluación) por la Universidad de Columbia de Nueva York. Catedrático de Métodos de Investigación y Diagnóstico en Educación en la Universidad del País Vasco.

bajar a nuestro ritmo, con objetivos cambiantes según vamos descubriendo cosas y planteándonos nuevos interrogantes. Nuestras reflexiones no se han limitado al contenido de los datos, sino que han sido también sobre la función de los distintos profesionales de la educación, los cambios de sus roles según legislaciones y prácticas nuevas y sobre problemas educativos y sociales con los que se relacionan los datos de que disponemos.

Trabajamos los sábados, cuando la universidad está tan quieta. Las horas se pasan volando y nos cuesta interrumpir los análisis y las reflexiones, pues la tentación siempre es hacer un análisis más que responderá a esa nueva pregunta tan interesante que los datos nos acaban de sugerir.

Aunque el centro sobre el que tenemos los datos no es excesivamente grande, y por lo mismo no hay demasiados alumnos por curso, en total tenemos datos de unos 1.500 alumnos y el fichero base de datos es de 675.000 bytes. Como uno de nosotros trabaja en ese centro tenemos muchos otros datos sobre los mismos alumnos, además de un conocimiento bastante completo de las características del contexto educativo y social.

Las preguntas que nos planteamos inicialmente eran bastante abiertas: ¿Qué nos dicen los datos sobre lo que hacemos, qué información nos aportan? ¿Sirven para algo? ¿Ese algo compensa los costes de obtenerlos? ¿Cómo les afectan los cambios legislativos y teóricos de la Psicología, Psicopedagogía y Pedagogía? Luego seguimos preguntándonos también hasta qué punto sirven para describir, diagnosticar y predecir en el presente, a corta y larga distancia. También si sirven para «liberar» de estereotipos castros a ciertos alumnos etiquetados con datos de ambientes siempre cerrados. Tenemos curiosidad también por ver hasta qué punto se confirma en nuestro caso lo que dicen los libros de texto y los gurus de moda en este momento. Los tres investigadores somos en general bastante escépticos respecto a las modas que van y vienen, y respecto a las legislaciones que sacralizan a los gurus con poder. La legislación que más nos interesa es la siguiente, la que va a venir, que siempre se espera que va a ser mejor que la presente. Recalcamos la importancia de las convicciones propias y del espíritu crítico, en parte porque hemos estado metidos en bastantes guerras en distintos ambientes, sobre todo en el de la Educación.

Los objetivos iniciales de la investigación eran, pues, abiertos y conforme los datos nos han ido hablando hemos ido planteándonos objetivos nuevos y/o adicionales, y hemos hecho los análisis correspondientes.

Estamos cerca del final de la investigación. Ya informaremos debidamente sobre la misma por medio de alguna publicación. Un factor de retraso puede ser que ahora sí pidamos financiación para lo que hemos hecho y para alguna ampliación menor que nos parece interesante.

Me he extendido bastante en contar algunas cosas de la investigación porque pienso referirme a este ejemplo algunas veces a lo largo del resto del artículo.

Las ideas que voy a exponer en las páginas siguientes sobre algunos retos metodológicos de la investigación las he estructurado en tres planos: el conceptual-epistemológico, el operativo-organizativo y el formativo-didáctico. El más importante objetivamente es el segundo, pues lo que interesa de verdad es que se investigue más y mejor. En el tercero estamos implicados mucho la mayoría de los que hablamos y

enseñamos metodología de la investigación. Sin embargo la mayor parte de la discusión del artículo va a ser en el plano conceptual-epistemológico, porque es el que afecta a los otros dos planos.

Antes de pasar a la parte principal del artículo voy a dejar claros algunos supuestos personales que son como el punto de partida desde el que planteo los retos y las posibles discusiones. Son precisiones iniciales que pueden servir para entender mejor lo propuesto en el resto del artículo, aunque ya sobre ellas quepan discusiones, pues varias son posicionamientos que más de uno no compartirá.

I. PUNTO DE PARTIDA

La perspectiva es de *Ciencias Sociales*, de no encerrarse en una disciplina de la Psicología, Pedagogía, Sociología, etc. sino trascender los límites de las distintas disciplinas, abrirse a otras perspectivas. El resultado será que mejoraremos la nuestra y que relativizaremos un poco lo que hacemos. Nos hará ver que ciertos posicionamientos que creemos superados continúan teniendo partidarios en otras disciplinas, cuyas razones haremos bien en atender. O nos obligará a plantearnos ciertas cuestiones que todavía no nos hemos planteado seriamente. Todo ello sin ignorar las diferencias reales que hay entre la investigación educativa, sociológica, psicológica o antropológica.

Se quiere evitar una *actitud* de beligerancia. Cada investigador es hasta cierto punto producto y víctima de una formación inicial más o menos adecuada y/o doctrinaria. Ciertos posicionamientos y formas de manifestarlos son más reflejo de características personales, sobre todo psicológicas, que de razones y convicciones intelectuales. Hay que respetar las distintas evoluciones y desde luego no caer en la trampa de las descalificaciones. La actitud de fondo en este artículo es de respeto a la inteligencia y competencia los otros profesionales. Como dice Coulter (1996, p. 1.226) la finalidad del quehacer científico no es ganar los debates a cualquier costa, sino descubrir si no verdades al menos una comprensión más profunda de los temas en cuestión.

Otro supuesto básico es ella *superación del debate cuantitativo-cualitativo* (Dendaluze, 1995). Esa referencia es la transcripción de la ponencia en el VII Seminario Nacional de Modelos de Investigación Educativa, organizado por la Asociación Interuniversitaria de Investigación Pedagógica Experimental (AIDIPE) y celebrada en Valencia en 1995. En esa referencia se explican los pasos seguidos para superar el debate. Varios de los más discutibles serán tratados posteriormente en el presente artículo. Esto hace inevitable el que haya ciertas repeticiones, pero la finalidad de la ponencia y de este artículo son distintas.

Otro punto de partida es el del *pluralismo integrador*, consistente en reconocer primero que se ha pasado de un exclusivismo a un pluralismo metodológico y en proponer después que ese pluralismo sea integrador de preocupaciones y soluciones. Lo cual supone en primer lugar un aprecio por la seriedad y profesionalidad de los que investigan de forma distinta de la propia; una voluntad de aprender unos de otros; y una disposición a flexibilizar los posicionamientos de forma que se admitan préstamos mutuos. Sobre las relaciones del pluralismo integrador con la inconmensurabilidad se hablará posteriormente.

Lo anterior no quiere decir que tengamos que renunciar a tener convicciones metodológicas básicas propias. Ni renunciar a hacer preferentemente un tipo de investigación que fundamentalmente es el más apropiado al tipo de problemas que más queremos estudiar.

No se trata tampoco de hacer combinaciones de dos metodologías, ni de síntesis, ni de sincretismos, cuyo producto es algo nuevo distinto de las dos cosas de las que partimos. Lo que se propone es una mejora de la metodología que más nos convence en general o que nos parece más apropiada para el problema que estamos estudiando. El producto con mentalidad de pluralismo integrador es una variedad de esa metodología, una variedad que asimila ciertos préstamos y que es tan legítima como la versión puritana de la metodología. Por ejemplo, si cualitativizamos la investigación cuasiexperimental (atención al proceso, participación de los agentes habituales de la intervención, etc.) no se trata de dejar de hacer investigación cuasiexperimental, sino de hacer una variedad de investigación cuasiexperimental (investigación cuasiexperimental cualitativizada), a veces más interesante que la investigación cuasiexperimental clásica.

La cuestión no es tampoco si integrar ciertas prestaciones de otras metodologías o profundizar en la metodología propia, sino de profundizar *e* integrar.

En ningún caso se trata de dictar a los otros investigadores, que trabajan con metodologías distintas de la nuestra, qué es lo que tienen que aprender de nosotros. La forma más útil de aprendizaje mutuo y de avance conjunto es que miremos nosotros qué nos puede ser útil de lo que hacen y dicen los otros. Es decir, no debemos decir «vosotros de nosotros tenéis que aprender...»; sino «a nosotros de lo vuestro nos vendría bien...».

Una última advertencia inicial se refiere a *tópicos y típicos*. Los alumnos siempre se sonríen cuando les advierto de los peligros del «TOP-TIP». Está relacionado con nuestra tendencia a las esquematizaciones didácticas, que con frecuencia son simplificaciones excesivas, caricaturas a veces, sobre todo de quienes no piensan como nosotros.

El «top» quiere decir que con frecuencia tenemos que aclarar que «tópicamente» un cuantitativista o un cualitativista actúa o razona así o asá. Y hay que añadir que «típicamente» actúa de esa forma; pero que hay muchos matices, variedades y diferencias entre unos cuantitativistas o cualitativistas y otros, que hay bastantes variedades y tradiciones de investigaciones cuantitativas y cualitativas. Por ejemplo, la investigación indicada al principio del artículo la mayoría la calificarían como cuantitativista (casi todos los datos y los análisis lo son). Pero el planteamiento del problema es abierto, flexible y emergente; que no son características que tópicamente explicamos como típicas de planteamientos cuantitativos.

2. PLANO CONCEPTUAL-EPISTEMOLÓGICO

En el plano conceptual-epistemológico vamos a discutir primero algunos retos sobre las líneas transversales de la investigación y después sobre cada uno de los aspectos críticos de la misma.

2.1. Líneas transversales

Hay que comenzar por discutir la idea misma de «líneas transversales» de la investigación. En un cuadro de doble entrada los aspectos críticos de la investigación (planteamiento del problema de investigación, planteamiento metodológico para la resolución del problema, etc.) pueden considerarse como la dimensión horizontal de la investigación y las líneas transversales como la dimensión vertical, como algo que afecta de arriba a abajo a los aspectos críticos. Aun en una conceptualización circular o en espiral del proceso de investigación, en cada momento de la misma se pueden distinguir esas coordenadas horizontal y vertical; es decir, analizar qué aspecto crítico estamos afrontando en cada momento y cómo le afectan las distintas líneas transversales.

En la primera presentación de la idea de líneas transversales en Valencia propuse tres: la psicosocial, la validez teórica y la objetividad-subjetividad. Es decir, tanto el aspecto crítico del planteamiento del problema de investigación como el del planteamiento metodológico para la resolución del problema, etc. se ven afectados por el posicionamiento respecto a la validez teórica, objetividad-subjetividad y consideraciones psicosociales.

La ampliación de la lista de líneas transversales supone que hay que discutir previamente al menos dos cuestiones: las bases conceptuales de la investigación y la relación de las líneas transversales con los paradigmas de investigación.

Las bases conceptuales de la investigación se refieren a los posicionamientos del investigador sobre formas de conocer, a sus preferencias por unas metodologías sobre otras y las razones de sus posicionamientos y preferencias en general y en investigaciones concretas en particular.

En la ponencia de Valencia se presentó lo referente a las bases conceptuales como el primero de los aspectos críticos que debe plantearse el investigador, anterior al planteamiento del problema de investigación, planteamiento metodológico, etc. Se dejaba la puerta abierta a considerar las bases conceptuales como una línea transversal. Las razones didácticas pueden ser un factor que aconsejen tratar las bases conceptuales como un primer aspecto crítico que se debe plantear el investigador antes de plantearse otros aspectos críticos.

Muchos traducirán el tipo de cuestiones tratados en las bases conceptuales a las tratadas en el debate y posicionamiento paradigmático, en que se deben resolver coherentemente cuestiones de axiología, ideología, ontología, epistemología, metodología y de técnicas. Las razones principales para no identificar las bases conceptuales con los paradigmas son dos. Una es el cansancio respecto al uso y abuso de los paradigmas. La otra razón es la rigidez habitual en la aplicación de los paradigmas a la problemática de la investigación.

En cuanto a la primera razón hay que reconocer que la idea de «paradigma» es una buena idea. De ahí su aceptación y popularidad. Tanta popularidad que se ha utilizado abusivamente, multiplicando sus sentidos. Sabido es que Kuhn tuvo cada vez más reservas respecto al término paradigma, que lo llegó a considerar como una condena a cargar hasta su muerte y que raramente lo utilizaba en los últimos años de su vida.

La solución que proponen algunos autores es la vuelta a la lógica y a la epistemología. Es decir, volver a plantearse las cuestiones básicas del criterio de verdad, de la lógica de la justificación, de la fuente del conocimiento, de su certeza, etc. (Kvale, 1996). Cada una de esas cuestiones se presta a distintas discusiones. Por ejemplo, Meana (1996) ha considerado al mismo Kuhn el guillotinator de la lógica de la investigación científica.

En cuanto a la segunda razón, la rigidez en la aplicación de los paradigmas de investigación, se trata de librarse del corsé que a veces se autoimponen algunos o que nos quieren imponer a otros. Esa rigidez se manifiesta en el esfuerzo por reducir el número de paradigmas en cuanto a la investigación a tres o cuatro, definirlos y categorizarlos excluyentemente, de forma que hay que optar por un paradigma y luego llevar la coherencia a no salirse del carril correspondiente desde la axiología hasta la última técnica asociada al paradigma elegido.

En toda esta temática de las bases conceptuales y/o paradigmas hay de entrada al menos dos peligros. El primero es el de separar excesivamente sobre todo ontología, epistemología y metodología (Kemmis, 1993). El segundo es el de unirlas demasiado estrechamente.

En cuanto al primer peligro hay que reconocer que una cierta axiología casa más y mejor con una cierta ideología, y ésta con ciertos fines de la investigación, con una ontología, epistemología, metodología y con ciertas técnicas.

En cuanto al segundo peligro la realidad es que es imposible y desaconsejable evitar prestaciones mutuas entre los paradigmas. Imposible porque no se pueden definir los paradigmas de forma exclusiva de arriba a abajo y de forma excluyente en todas las dimensiones del debate paradigmático. Desaconsejable porque es condenarse a un empobrecimiento puritano. Tener en cuenta algunas de las cuestiones y soluciones que se plantean los investigadores de otros paradigmas mejora la forma propia de investigar.

Ambos peligros pueden perpetuarse por lo cerrado que hacemos el ambiente de cada disciplina o de cada bandería. Ciertas plantas crecen mejor en ciertos ambientes y son las únicas aceptadas por el aire cargado que tendemos a crear cada comunidad científica.

No se puede seguir adelante sin afrontar la cuestión de la inconmensurabilidad, a distinguir de la incompatibilidad, que ha sido tratada menos sistemáticamente. De La Orden y Mafokozi han tratado de la inconmensurabilidad en un artículo reciente (1997). Por un lado resaltan las razones por las que nos planteamos la inconmensurabilidad (vocabulario, gramática, problemas, perspectivas, interpretaciones, soluciones asumibles, etc.) y por otro discuten algunas objeciones que precisan los límites de la inconmensurabilidad. Cuando se aplica ese concepto a las metodologías de la investigación la postura de los autores es contraria a las rigideces y a los exclusivismos.

El posicionamiento personal sobre la cuestión de la inconmensurabilidad es que se da a ciertos niveles, sobre todo filosóficos, en el que algunas ideas son claramente inconmensurables; por ejemplo, entre ciertos posicionamientos fenomenológicos y hermeneúticos. Pero hay que preguntar a los partidarios de recalcar la inconmensura-

bilidad si la llevan a todos los niveles. Se puede objetar a esa postura yendo de menos a más, comenzando por los niveles técnicos y acabando en los ontológicos. ¿No sumar 2+2 porque eso es cuantitativo y yo soy cualitativista? ¿No permitirse ciertas formas de recogidas de datos «propias» de otro paradigma? ¿Ninguna forma de triangulación? ¿Los cuantitativistas deben renunciar a toda forma de verificación de las conclusiones con los participantes? ¿Los cualitativistas deben excluir las explicaciones alternativas con un cierto rigor, incluso recogiendo datos sobre factores que luego pueden ser fuentes de explicaciones alternativas? Son sólo ejemplos para decir que evitar prestaciones de unas formas de investigar a otras es ilógico y empobrecedor.

De hecho es difícil encontrar investigadores que oficialmente se catalogan en una forma de investigar que no acepten en la práctica prestaciones de otras formas de investigar. Dejemos las inconmensurabilidades para los niveles más altos, el de la ontología y/o epistemología, para los filósofos, que gozan más de esas inconmensurabilidades y de otras incompatibilidades. El pluralismo integrador apunta a soluciones más personales del investigador concreto que es quien determina la coherencia de sus bases conceptuales y fija los límites de las prestaciones que acepta para acercarse mejor a la realidad. Es decir, en cuanto a la inconmensurabilidad los investigadores en Ciencias Sociales nos tendremos que preguntar inconmensurabilidad de qué posicionamiento con cuál otro, a qué nivel y en cuanto a qué.

Según cuál sea el resultado de la discusión anterior sobre paradigmas y/o epistemología, inconmensurabilidad y/o pluralismo integrador, etc. se puede plantear la ampliación de la lista de líneas transversales. No han aparecido razones para eliminar ninguna de las tres propuestas inicialmente en Valencia (psicosocial, validez teórica, objetividad-subjetividad). En la misma línea serían primeras candidatas a la ampliación de la lista la causalidad (con posicionamientos sobre las cuestiones de indeterminación e impredecibilidad) y el cambio.

Antes de discutir el significado concreto de esas y otras posibles candidaturas, los que somos partidarios de centrar más nuestros retos y discusiones en las dimensiones metodológicas y técnicas de la investigación deberíamos explorar la idea de convertir ciertos aspectos planteados en el debate paradigmático en líneas transversales. Así pasarían a ser líneas transversales, por ejemplo, la axiología y la ideología que legítimamente podrían influir o no en el planteamiento del problema de investigación, en el planteamiento metodológico, etc. La cuestión de los grandes fines de la investigación (explicar, comprender, cambiar, liberar, etc.) se pueden convertir también en líneas transversales. Estos grandes fines hay que distinguirlos claramente de los objetivos próximos «de» la investigación como investigación. Los objetivos conseguidos «por» realizar una buena investigación, estarían más unidos a los grandes fines. De la ontología y epistemología también se podrían deducir algunas líneas transversales; por ejemplo, del criterio de verdad, según sea el dato externo o interno. Esto debería traducirse con rigor, entre otras cosas, al momento y modo de los análisis y de sacar conclusiones.

Éstas y otras ideas sobre las líneas transversales necesitan una mayor discusión y operativización. En todo caso no se trata de prescindir totalmente de la idea de paradigmas y sí de salvar las preocupaciones que se reflejan en los distintos paradigmas.

2.2. Aspectos críticos

Hablamos de aspectos críticos en un doble sentido: primero, en cuanto que son los aspectos, entre las muchas cosas que hacemos los investigadores al investigar, en los que más nos jugamos el que nos salga bien la investigación; y segundo, en cuanto que son en los que más nos fijamos a la hora de criticar/valorar una investigación. Reconociendo siempre que una cadena es tan fuerte como su eslabón más débil; es decir, que una investigación puede fallar por cualquier aspecto no bien resuelto.

Se pueden plantear dos cuestiones, una formal y otra concreta. La cuestión formal es si realmente en cuanto a metodologías de la investigación hay unos cuantos aspectos básicos que compartimos todos los investigadores empíricos, tanto cuantitativos como cualitativos. La cuestión concreta es intentar hacer una lista de aspectos posiblemente comunes y ver cuántos los aceptamos desde las distintas metodologías.

En Valencia propuse cinco «aspectos críticos» de toda investigación empírica: bases conceptuales, planteamiento del problema de investigación, planteamiento metodológico para la resolución del problema, datos (sujetos, datación, análisis) y globalización (con conclusiones). Ninguno es novedoso. Ni en general, ni para las metodologías cuantitativas, ni para la mayoría de las tradiciones cualitativas.

Más recientemente, en las clases de metodología suelo hablar de 5^{1/2} aspectos críticos. El medio se debe a la discusión de si los informes de la investigación son un aspecto crítico o no. Personalmente me inclino a que sí, sobre todo por la importancia metodológica que tienen en algunas metodologías cualitativas los informes que se hacen en distintos momentos de la investigación.

Conviene resaltar que los aspectos críticos no se tratan como estadios o fases del proceso de investigación. Hacerlo puede reforzar la idea de la linealidad del proceso de investigación. Es decir, no se quiere dar la imagen de que investigar es primero plantearnos el problema de investigación, luego la metodología, luego recoger datos, etc. Lo normal es que, por ejemplo, al hacer los análisis tengamos que volver a los problemas de investigación; o al pensar en recoger los datos repensemos el diseño, etc. Este dinamismo entre los distintos aspectos críticos de la investigación, con continuas marchas atrás y adelante de unos a otros, se debe dar tanto en las investigaciones cuantitativas como cualitativas. Así mismo no va contra la idea de aspectos críticos el que en una investigación concreta se pase sucesivamente por varios problemas, diseños, datos, etc. En este caso el investigador sí que tiene que tener claro en cada momento concreto qué aspecto crítico está tratando de resolver.

Al hablar de las líneas transversales ya se ha apuntado la posibilidad de considerar las *bases conceptuales* o como un aspecto crítico o como una línea transversal, y se han tratado en el apartado de las líneas transversales. Aquí vamos a centrarnos en los otros cuatro aspectos críticos. Del medio aspecto crítico, los *informes*, valga en este artículo la mención ya hecha al reto que se plantea en torno a los mismos. Con mentalidad de pluralismo integrador las investigaciones cuantitativas ganarían bastante si se tomaran los informes durante el proceso de investigación tan en serio como lo hacen los

cualitativistas. Sería cuestión de utilizar más y mejor los informes de progreso que ya se hacen en bastantes investigaciones cuantitativas.

Centrándonos ya en el aspecto crítico del *planteamiento del problema de investigación* el núcleo del mismo está en resolver bien lo referente a los objetivos de la misma. Tanto si la investigación va a limitarse a resolver un problema con unos objetivos bien pensados y fijados como si los objetivos iniciales se piensan como punto de partida que seguramente nos llevarán a otros.

Se supone que en la gran mayoría de los casos los objetivos estarán bien contextualizados y/o fundamentados en una buena revisión de la literatura; que, según gustos y convicciones, se plantearán o no un modelo teórico y unas hipótesis. Por ejemplo, Ashworth (1996, 287) nos previene sobre el peligro de importar inadvertidamente teorías y métodos que distorsionarían la descripción del mundo real. Para él es un principio metodológico importante que el investigador tiene que comenzar por poner a un lado los supuestos sobre la naturaleza de fenómeno que estudia.

Las diferencias entre las distintas metodologías cuantitativas y cualitativas se centran tanto en cuanto al qué investigan como al cómo plantean los problemas. Tanto unas diferencias como otras hay que clarificarlas más, dejando al lado ciertos tópicos. Por ejemplo, el de que los problemas de las investigaciones cualitativas, en contraposición a los de las cuantitativas, son flexibles, abiertos y emergentes, y que los problemas de las investigaciones cuantitativas no. Para contrarrestar este tópico me remito al ejemplo de nuestra investigación presentada al principio del artículo.

Volviendo a la cuestión central de los objetivos, no se excluye que haya buenas investigaciones sin objetivos explicitados. Habrá que determinar cuándo esta forma de proceder tiene mayor valor heurístico. Puede ser un buen proceder que no justifica la pereza de quienes antes de moverse no se toman la molestia de pensar lo que les vendría bien pensar, de contextualizar, de explicitar ciertas dudas y preguntas germinales, etc.

También se puede discutir si es bueno distinguir entre los grandes fines de la investigación y los objetivos próximos de la investigación. Metodológicamente interesan más éstos, por muy influidos que estén por los grandes fines axiológicos, ideológicos o paradigmáticos. Se propone que con unos mismos fines axiológicos, etc. podemos llegar a investigaciones concretas con distintos objetivos próximos.

Es de esperar que provocará menos discusiones la idea de que sobre un mismo tema (integración, maltrato, bilingüismo, etc.) cabe plantear distintas investigaciones cuantitativas o cualitativas, según gustos y convicciones del investigador, todas metodológicamente bien planteadas.

El tercer aspecto crítico sobre el que tendremos que clarificar cuestiones y soluciones es el del *planteamiento metodológico* para la resolución del problema de investigación. Son parte de ese planteamiento tanto lo referente al diseño/plan de la investigación como la determinación de en qué sujetos/elementos se va a realizar la investigación, qué datos se van a buscar y cómo se van a analizar. En Valencia se propuso estudiar como tercer aspecto crítico lo referente al diseño/plan de la investigación y como cuarto lo referente a los datos (sujetos, datación, análisis).

Por suerte la mayoría de los investigadores no se dejan atrapar en las diferencias terminológicas en cuanto a diseño, plan, esquema, estrategia, etc. Tampoco pasa gran cosa si una vez más cada uno tenemos preferencias terminológicas con tal de que definamos los términos que empleamos para que nos entendamos traduciéndonos mutuamente.

Las definiciones del diseño/plan de la investigación seguirán siendo distintas por ser distintas las concepciones de la investigación; pero en todas ellas habrá un componente de previsión-organización de las acciones del investigador, otro de respuesta a los interrogantes-objetivos de la investigación y otro de conseguir la calidad (validez/credibilidad) de la investigación.

Habría que discutir por qué en bastantes proyectos e informes de investigación no se trata expresamente de su diseño/plan. ¿Se da por supuesto que no hay alternativas al que se emplea, que ése es necesariamente el adecuado, que no hay que razonar la opción que se adopta? Suele pasar en los colectivos de investigadores encerrados en sí mismos.

El investigador que por la naturaleza de su investigación busque la validez de la misma encontrará que los conceptos, amenazas y técnicas básicas están bastante estandarizados. Tal vez habría que ampliar más formalmente y expresamente su lógica a la investigación cuantitativa no causal. Además algunos pensamos que la validez de constructo está bien, pero que la validez teórica (Dendaluze, 1997) es más englobadora y útil. Habría que discutirlo más.

Quienes prefieren los criterios de credibilidad, transferibilidad, dependencia y confirmabilidad (Colás, 1995) pueden plantearse si los que propone Kvale (1996) de coherencia, defensibilidad, verdad pragmática y verdad comunicativa no deben ser atendidos también y encontrar alguna formulación englobadora simplificada. Los procedimientos cualitativos para conseguir la calidad de la investigación, comoquiera que finalmente se categoricen y definan los criterios, parecen estar en claro camino de estandarización.

Parece innegable que ha habido prestaciones mutuas entre las metodologías cuantitativas y cualitativas en cuanto a este aspecto crítico. Puede pensarse el paralelismo, por ejemplo, entre generalizabilidad y transferibilidad, objetividad y neutralidad, etc. Aún así habrá que trabajar más para estructurar mejor cómo van a asimilar las metodologías cuantitativas la legítima preocupación por el contexto, la globalidad y la colaboración de los participantes. Viceversa las metodologías cualitativas la preocupación por las explicaciones/hipótesis alternativas, con atención a los factores que pueden dar origen a esas explicaciones alternativas; por ejemplo, asegurando que se va a disponer de datos sobre esos factores.

Una de las justificadas víctimas del merecido arrinconamiento del positivismo (que no se debe identificar con el cuantitativismo) en las Ciencias Sociales ha sido la definición constitutiva y operativa de las variables. Green y Powell (1990) hablan del fallo en primer lugar del intento de positivistas y positivistas lógicas de dar una definición empírica a los significados, fallo reconocido según ellos por el mismo Carnap; fallo en segundo lugar del operacionismo, aliviado de su miseria por su mismo originador,

Brigidman (1978) y sepultado por Koch (1959); aunque son conscientes de que todavía hay gente que mantiene y defiende el operacionismo.

Según lo anterior está bien que no definamos las variables constitutiva y operativamente. ¿Pero quiere decir eso que está bien que no las definamos claramente de ninguna forma? Los términos, constructos, variables, factores, etc. que utilizamos en las Ciencias Sociales suelen responder a distintas conceptualizaciones o tener matices distintos según quién los usa. Parece obligado por ello que el investigador defina el sentido en que él los usa. Puede no ser suficiente una sola definición, sólo la conceptual o solo la concreta, de los términos que utilizamos. A veces pueden venir bien las dos definiciones, conceptual y concreta. Si algunos piensan que este proceder es dejar entrar por la ventana a las definiciones constitutivas y operativas que expulsamos por la puerta, habrá que discutirlo. Aquí se defiende que una buena definición de las variables contribuye por supuesto a la validez teórica de la investigación, también a la validez interna sobre todo por parte de la datación, sin excluir la validez interna por parte del diseño y del análisis. Y que la definición de las variables es también necesaria en las investigaciones cualitativas por razones análogas.

El aspecto crítico de los *datos* hay que tratarlo en sus tres subaspectos: sujetos, datación y análisis. En cuanto al primero, sujetos de la investigación, parece que las ideas están claras y que hay poco que debatir. Es decir, se determina quiénes son los sujetos o elementos que se quieren investigar, se decide si va a ser toda la población o una muestra, y en este caso de qué tamaño va a ser ésta y cómo se va a seleccionar. Las técnicas aleatorias y no aleatorias están bien explicadas y, por lo mismo, es cuestión de opciones del investigador qué técnicas utilizar según el problema que esté investigando.

Quedan problemas «menores» como superar ciertos tópicos que asocian los grandes números y la finalidad nomotética con la investigación cuantitativa. ¿Tiene, por ejemplo, finalidad nomotética la investigación mencionada al principio del artículo? Por otra parte los investigadores cuantitativistas harán bien en plantearse seriamente la buena utilización de los informantes-clave (Rodríguez y Otros, 1996, 135-139).

Son evidentes las implicaciones de las decisiones referentes a los sujetos en cuanto a los sesgos de selección y para la generalizabilidad/transferibilidad. Lo menos que se puede pedir a todos los investigadores es coherencia entre las distintas elecciones que hagan y una cierta prudencia a la hora de aplicar las conclusiones a otras poblaciones, momentos y situaciones.

El término *datación* engloba tanto la recogida como la elaboración de los datos, tanto cuantitativos como cualitativos. Es un término expresivo y breve para el que hasta ahora los que lo utilizamos no hemos encontrado una alternativa mejor.

La recogida de datos es un aspecto bastante y bien trabajado por las distintas metodologías cuantitativas y cualitativas. En cuanto a la elaboración de tipos y niveles de datación debería hacerse más por parte de las metodologías cualitativas. En este caso no basta con distinguir entre los datos recogidos los directamente numéricos y los verbales; y entre éstos los convertibles a números (a escalas nominales, ordinales y, por qué no, también a escalas de intervalo) y los que se mantienen como verbales. Por de pronto puede haber información útil en todos ellos y se deben analizar todos. La elaboración de datos que hay que trabajar más es la referente a datos verbales que se que-

dan en verbales, ya que también en ellos hay clases de datos que seguramente se prestan a análisis distintos.

La amplia aceptación teórica y sobre todo práctica de la triangulación por parte de casi todas las metodologías ha constituido un gran avance que ha servido no solo para contactar mejor con la realidad, sino también para romper muchas barreras rígidas entre ciertas metodologías. Aún así hay resistencias como las que expresa Ashworth (1996), quien encuentra a Kvale (1996) tal vez demasiado tolerante al combinar entrevistas y cuestionarios. En su caso la inconmensurabilidad llega al nivel técnico. También Fernández Cano (1995, 151) se muestra contrario a triangulaciones injustificadas por responder a supuestos ontológicos y epistemológicos distintos.

Quienes estamos abiertos a las prestaciones mutuas entre metodologías cuantitativas y cualitativas las posibilidades son grandes. Además de la triangulación, están los conceptos de validez y fiabilidad, la concepción múltiple de la realidad, la importancia de los datos «desde dentro», los estudios en profundidad, la importancia de los datos obtenidos en situaciones naturales y enmarcados en una globalidad, etc. Habrá que discutir estas y otras posibilidades, sus límites y limitaciones.

En cuanto al *análisis de datos* la situación ha sido muy distinta en el campo cuantitativo y cualitativo. Se lleva mucho tiempo utilizando y mejorando la Estadística y el investigador que quiere hacer análisis estadísticos tiene a su disposición muchas opciones e instrumentos muy preparados, a pesar del debate en torno a la prueba de hipótesis (Wilcox, 1998; Shrout, 1997; Glass, 1995; Cohen, 1994; Schmidt, 1994; Meehl, 1990). El progreso en el campo cualitativo ha sido muy notable en poco tiempo y cada vez los procedimientos son más adecuados, sofisticados y con instrumentación informática.

Aún así queda mucho margen para retos y debates. El primero tan elemental como el de la mayor o menor separación de los análisis cuantitativos y cualitativos. La gran mayoría de los cualitativos (Rodríguez y Otros, 1996; Colás, 1994) no renuncian a las aportaciones de los análisis estadísticos, por muy en su sitio que haya que ponerlos. Muchos cuantitativos sólo se molestan en analizar los datos verbales si se cuantitativizan al menos en forma de categorizaciones. Entre cuantitativos se da con frecuencia una confusión en cuanto al término mismo de datos «cualitativos», que para muchos son los datos categorizados a un nivel nominal u ordinal. Es un sentido habitual, pero hay otros sentidos.

Los investigadores cuantitativos no pueden dejar sin analizar información relevante y rica simplemente porque los datos son verbales y no se puede utilizar la Estadística o no se puede utilizar fácilmente. ¿Por qué si los cualitativos no renuncian a la Estadística los cuantitativos no se abren igualmente a la posibilidad de análisis no estadísticos?

Además bastantes cualitativos podrían plantearse el adaptar la buena lógica de los análisis estadísticos en cuanto a no aceptar cualquier diferencia como significativa (aunque se emplee el término no en sentido estadístico), o como prueba de que los datos confirman sus expectativas. Se oye decir con frecuencia que gracias a tal programa de intervención el colectivo X está mejor que antes. Es legítimo preguntar a quien así habla si la determinación del mejor o peor es a «ojímetro», si es poco o mucho

mejor, si hay datos para comparar el antes y después, si es gracias al programa o a otras cosas, si hay datos para clarificar esto último, etc.

Respecto al aspecto crítico de la *globalización* sería bueno otra vez no enredarse en distinciones terminológicas entre obtención de resultados en los análisis, interpretación de los resultados, discusión de los mismos, integración de la investigación, *globalización*, etc. Cada término resalta algo y cada matización encierra un enriquecimiento. Pero en lo que se quiere insistir en este aspecto que llamamos *globalización* y que además consideramos crítico es en que el investigador debe levantar la vista de cada interrogante, objetivo, análisis, resultado y conclusión particular hacia el conjunto de la investigación, para tratar de ver lo que significan y aportan en conjunto todas las cosas particulares que se han hecho; qué se ha aprendido en el conjunto de la investigación.

Las conclusiones de una investigación no son sólo la suma de las conclusiones particulares que se han ido obteniendo en distintos momentos particulares de la investigación o sobre interrogantes y objetivos particulares. Hay que sacar también conclusiones basadas en la interrelación de las conclusiones particulares. Aparte de que es legítimo y casi una obligación el mirar también cómo se relaciona lo que hemos visto en nuestra investigación con las teorías y autores que han tratado el tema y con las investigaciones previas a la nuestra. De aquí surgen también algunas conclusiones.

En el caso de algunas investigaciones cualitativas, las más tópicamente típicas, es el momento de «reconstruir» objetivos y conclusiones parciales que han ido dándose a lo largo de todo el proceso de investigación, reconstruirlas en un todo más global y globalizador o al menos agruparlas al final de la investigación. Lo pueden hacer por sí mismos los lectores del informe de la investigación; pero quien mejor puede hacerlo es el mismo investigador.

Por supuesto, según nos recuerdan Huck y Sandler (1979, 239-244) la interpretación de los resultados es distinta según quién la haga. Habrá quien aplique también aquí el concepto de inconmensurabilidad de Feyerabend (1981). Decía que las interpretaciones científicas dependen del marco teórico y por lo mismo no pueden compararse hipótesis rivales, pues surgen de supuestos teóricos distintos. Esta discusión puede llevar a situar y operacionalizar mejor el concepto de «acuerdo entre interpretadores» desde la ontología y epistemología hasta la metodología y las técnicas.

Otro punto de debate y posible clarificación es hasta qué punto las conclusiones deben estar atadas a los datos o hasta dónde pueden llegar las inferencias interpretativas del investigador. Cronbach (1987, 22) nos recuerda que todas las conclusiones son inferencias, pues van más allá de los datos con la ayuda de supuestos y presuposiciones. Es decir, las conclusiones no se basan únicamente en los datos, sino también en una amplia base de supuestos, muchos de ellos susceptibles de discusión.

3. PLANO OPERATIVO-ORGANIZATIVO

Puesto que el buen quehacer estará muy influido por consideraciones del plano conceptual-epistemológico, conviene que antes de hacer nada cada investigador se aclare sobre sus propias ideas y convicciones sobre la investigación. Lo que cierta-

mente siempre debe tener claro es qué clase de investigación está haciendo, y por qué esa y no otra.

En concreto si defendemos el pluralismo integrador se nos amplían enormemente las posibilidades; pero también nos obliga a conocer mejor más alternativas antes de adoptarlas en la práctica. Esto normalmente nos lleva a trabajar más en equipos de investigación más interdisciplinares y con distintos saberes metodológicos y técnicos.

Uno de los retos es superar la organización legislativa en áreas de conocimiento y departamentos que ha reforzado la tendencia a encerrarnos en guetos profesionales y como consecuencia en conceptualizaciones y metodologías dominantes en nuestros ghettos. Tendremos que hacer un esfuerzo por escucharnos más, discutir más, y estar dispuestos a aprender unos de otros.

Tendremos también que operacionalizar mejor que hasta ahora los intercambios, discusiones y asesoramientos varios por correo electrónico e Internet. Lo hábil será que a partir de un cierto volumen de trabajo lo convirtamos en parte de nuestro trabajo contabilizado y remunerado.

Un «detalle» muy práctico es el de las líneas de investigación en general y en concreto de los metodólogos. Habrá que trabajar seriamente, por bien que atendamos a la mejora del método y a su docencia, para que todo metodólogo tenga al menos una línea de investigación. Esto tiene muchas ventajas, pues entre otras cosas enraiza el saber metodológico en un campo concreto posiblemente con mayor eco social.

4. PLANO FORMATIVO-DIDÁCTICO

Los retos en este plano formativo-didáctico tienen dos frentes, el formativo por un lado y el didáctico por otro. El formativo incluye la formación inicial y la permanente. La formación inicial se consigue normalmente como parte de los estudios universitarios conducentes al título que sea. Las desigualdades entre las distintas titulaciones en Ciencias Sociales son muy grandes tanto en cuanto a las asignaturas metodológicas obligatorias como en cuanto a las posibilidades de formación que ofrecen en forma de asignaturas optativas, practicums/práctica, existencia o no de un itinerario de investigación-evaluación, configuración del tercer ciclo y programas de doctorado, desarrollo de tesis, inclusión de alumnos en equipos de investigación, etc.

A partir de la formación metodológica inicial que tenga, cada profesional interesado en hacer investigación se tiene que plantear su formación permanente. Se sabe que a investigar se aprende sobre todo haciendo investigación, preguntando mucho, estudiando lo que haga falta y no aceptando como definitivos los límites y las lagunas que uno tenga. En cada investigación en que uno participe debe forzar sus propios límites. Oportunidades caras y baratas de formarse siempre hay, comenzando por aprovechar las ofertas ordinarias de los planes de estudio del entorno en que uno se mueve. También jugarán un papel fundamental los asesoramientos y pequeñas clases concretas por parte de profesionales que saben más que uno mismo sobre ciertos métodos y técnicas de investigación. La forma más directa de conseguir esta formación continua es entrar en equipos de investigación, cuanto más interdisciplinares mejor. Es condición estar dispuesto a aprender y a veces dedicar el tiempo necesario a estudiar.

Otra cuestión a plantearse es si escogemos ser investigadores de un solo método, por saber y por utilización real. Otra cosa es que queramos saber sobre todo de un método y sobre ciertas técnicas. Habrá que ver si aceptamos lo que Seidman (1993, 6) dice sobre los alumnos que hacen investigación cualitativa sin aprender los puntos fuertes de la investigación cuantitativa. Según él tienden a no entender ni la historia del método que están usando, ni las críticas del positivismo y de la experimentación de las cuales han surgido algunos enfoques de la metodología cualitativa. El se ha decantado por las metodologías cualitativas tras hacer una tesis doctoral cuantitativa, experiencia que no le gustó.

Se puede plantear por analogía si quienes se decantan casi exclusivamente por métodos y técnicas cuantitativas se pueden permitir en los tiempos actuales el no aprender los puntos fuertes de las metodologías cualitativas.

En la práctica muchas actitudes de aprendizaje están centradas en torno a la Estadística. Renunciar totalmente a ella es mucho renunciar. El problema concreto es cuánto saber de Estadística en todo caso y sobre todo si se hace preferentemente o exclusivamente investigación cualitativa. La respuesta general es que cuanto más Estadística se sepa mejor. Por comprender otras aportaciones y por tener abierta la opción de utilizarla si ha lugar, que normalmente la habrá. Es malo renunciar a una opción porque no se está en condiciones de saber utilizarla. Desde luego es del género tonto denigrar algo en el fondo porque no se sabe sobre ello. Es peor que el caso de las uvas verdes de la zorra de la fábula.

El argumento anterior es totalmente extensible a los cuantitativistas en cuanto a los análisis de datos verbales, que contienen tanta información y que son productos de técnicas (observación participante, entrevistas, documentos, etc.) que un investigador debe conocer al menos en sus grandes líneas.

El frente didáctico también nos presenta varios retos: Qué enseñamos sobre la investigación, cuántas metodologías y técnicas enseñamos, qué metodologías y técnicas enseñamos más, etc. Esto está condicionado por el entorno en que se desarrolla la función docente-didáctica y que es muy desigual en cuanto a posibilidades y medios. Comenzando por los planes de estudio de cada universidad, por el poder de los distintos departamentos y por el reparto de cargas docentes, etc. Las opciones son varias y por lo mismo caben varios debates y discusiones.

El primer debate es cuál es la mejor concatenación entre asignaturas de iniciación, profundización y especialización de métodos y técnicas de investigación, evitando repeticiones innecesarias de contenidos al mismo nivel de explicación. Después cabe discutir si centrarse o no casi exclusivamente en un tipo de metodología y de técnicas cuantitativas o cualitativas.

Con mentalidad de pluralismo integrador y como un punto de partida para una posible clarificación de ideas se propone que en los cursos de iniciación a la metodología de la investigación, incluso en los cursos de profundización, se estructure la docencia en torno a los aspectos críticos, explicando cuál es la esencia del aspecto crítico que se estudia; qué cuestiones se plantean y cómo las resuelven las distintas metodologías cuantitativas y cualitativas; cómo le afectan las líneas transversales y, si se quiere, exteriorizar qué podrían aprender los unos de los otros, las posibles prestaciones mutuas.

Una ventaja muy importante del enfoque del pluralismo integrador es la libertad con que se estudian los problemas metodológicos y la receptividad con que se escuchan todas las opiniones y aportaciones. Incluso un problema interesante que menciona Stallings (1995) adquiere una nueva dimensión. El habla de primera y segunda generación de profesores de metodologías cualitativas. Con la perspectiva del pluralismo integrador uno puede venir de donde venga, estar marcado por ciertas influencias y tener ciertas querencias metodológicas básicas. Eso no es obstáculo para una buena docencia con tal de que los alumnos oigan los posicionamientos y argumentos principales de los distintos enfoques, se fomente una actitud abierta a aprender de quien tenga algo que aportar, se favorezca la crítica de los puntos fuertes y débiles de las distintas metodologías y técnicas, y tengan la oportunidad de profundizar en aquellas concretas que más les atraigan.

Las asignaturas de especialización parece claro que deben centrarse en métodos o técnicas específicas, cuantitativas o cualitativas, a elección según el gusto del alumno. La habilidad de la oferta docente consistirá en conjugar las optativas que se ofrecen en las distintas titulaciones de forma que el alumno en cada cuatrimestre tenga realmente opción a escoger entre metodologías y técnicas cuantitativas y cualitativas.

CONCLUSIÓN

Los retos propuestos en ningún caso se presentan como una lista exhaustiva. Además en vez de éstos podrían haberse propuesto otros. Si aparecen otras listas, cada uno podrá escoger los que más le interesan o preocupan. Con el intercambio y de la discusión ganaremos todos. Es una tarea siempre inacabada que no terminará, como no terminarán otras cosas, en el año dos mil.

BIBLIOGRAFÍA

- Ashworth, P. (1996). The primacy of conversation. *The humanistic Psychology*, 24(2), 283-290.
- Bartolomé, M. (1990). *Elaboración y análisis de datos cualitativos aplicados a la investigación-acción*. Barcelona: PPU.
- Brewer, J.; Hunter, A. (1990). *Multimethod research*. London: Sage.
- Bridgeman, P.N. (1978). Operationism. In Hardy, W.G. (Ed.). *Language, thought and experience*. Baltimore, MD: University Park Press.
- Code, L. (1991). *What can she know? Feminist theory and the construction of knowledge*. Ithaca, NY: Cornell University Press.
- Cohen, J. (1994). The earth is round ($p < .5$). *American psychologist*, 49(12), 997-1.003.
- Colás, P. (1995). La metodología cualitativa. En P. Colás, L. Buendía: *Investigación educativa* (pp. 249-290). Sevilla: Alfar.
- Coulter, X. (1996). On rebuttals by scientists. *Contemporary Psychology*, 41(12), 1.226.
- Cronbach, L.J. (1987). Issues in planning evaluations. In R. Murphy; H. Torrance (Eds.). *Evaluating Education: Issues and methods* (pp. 4-35). London: Harper & Row.

- Dendaluze, I. (1995). Avances en los métodos de investigación educativa en la intervención psicopedagógica. *Revista de investigación educativa*, 26, 9-32.
- Dendaluze, I. (1997). La validez teórica de las investigaciones empíricas en Ciencias Sociales. *Revista internacional de estudios vascos*, 42(1), 77-101.
- Fernández Cano, A. (1995). *Métodos para evaluar la investigación en Psicopedagogía*. Madrid: Síntesis.
- Feyerabend, P.K. (1991). *Tratado contra el método*. Madrid: Tecnos.
- Glass, G.V. (1995). Review of Harris Cooper. Hedges, L.V. (Eds.). *The handbook of research synthesis*. New York: Russell Sage Foundation. *Contemporary Psychology*, 40(8), 736-738.
- Green, C.D., Powell, R. (1990). Comment on Kimble's generalism. *American psychologist*, 45(4), 556-557.
- Guba, E.G., Lincoln, Y.S. (1989). *Fourth generation evaluation*. Newbury Park, California: Sage.
- Huck, S., Sandler, H.M. (1979). *Rival hypotheses: Alternative interpretation of data based conclusions*. New York: Harper.
- Kemmis, S. (1993). Foucault, Habermas and evaluation. *Curriculum studies*, 1(1), 35-54.
- Koch, S. (Ed.) (1959). *Psychology: The study of a science*. New York: McGraw-Hill.
- Kuhn, T.S. (1962). *La estructura de las revoluciones científicas*. Madrid: Fondo de cultura económica.
- Kvale, S. (1996). *InterViews: An introduction to qualitative research interviewing*. Thousand Oaks, CA: Sage.
- Meana, L. (1996). La minirrevolución de Thomas Kuhn. *El País-Babelia*, 13-7-96, p. 14.
- Meehl, P.E. (1990). Why summaries of research on psychological theories are often uninterpretable. *Psychological reports*, 66, 195-244.
- Miguel Díaz, M. de (1988). Paradigmas de la investigación educativa española. En I. Dendaluze (Coord.). *Aspectos metodológicos de la investigación educativa*. Madrid: Narcea, pp. 60-77.
- Munarriz, B. (Coord.) (1997). *Investigación participativa*. Donostia: Ibaeta Pedagogía.
- Orden, A. de la; Mafokozi, Joseph (1997). Implicaciones de algunos planteamientos epistemológicos post-positivistas en la investigación educativa. *Bordón* 49(4), 347-358.
- Rodríguez, G., Gil, J., García, E. (1996). *Metodología de la investigación cualitativa*. Archidona (Málaga): Aljibe.
- Schmidt, F.L. (1994). Quantitative methods and cumulative knowledge in Psychology: Implications for the training of researchers. *The Score*, 17(4), 5-6.
- Searle, J.R. (1997). *La construcción de la realidad social*. Barcelona: Paidós.
- Seidman, I.E. (1991). *Interviewing as qualitative research: A guide for researchers in Education and the Social Sciences*. New York: Teachers College Press.
- Shadish, W.R. (1989). Critical multiplism: A research strategy and its attendant tactics. In Sechrest, L., Freeman, H., Mulley, A. (Eds.). *Health services research: A focus on AIDS*, pp. 5-28. Rockville, MD: Agencia Health Care Police Res.
- Shrout, P.E. (1997). Should significance tests be banned? Introduction to a special section exploring the pros and cons. *Psychological science*, 8, 1-2.

- Stallings, W.M. (1995). Confessions of a quantitative educational researcher trying to teach qualitative research. *Educational researcher*, 24(3), 31-32.
- Wilcox, R. (1998). How many discoveries have been lost by ignoring modern statistical methods? *American psychologist*, 53(3), 300-314.

TRABAJOS DE INVESTIGACIÓN

ELABORACIÓN DE UN SISTEMA DE CATEGORÍAS PARA EL ANÁLISIS SOCIOMÉTRICO: SU APLICACIÓN EN AULAS MULTICULTURALES

Flor Cabrera, M.º Ángeles Marín, Julia V. Espín y Mercedes Rodríguez
Universidad de Barcelona

RESUMEN

En este artículo se presenta un sistema de categorías para analizar la situación psicosocial de las personas dentro de un grupo. Supone una forma de análisis de la información más cualitativa y dinámica que la que ofrecen los tradicionales índices sociométricos. En las categorías se contempla no sólo la cantidad de elecciones, rechazos y percepciones emitidas y recibidas, sino también la calidad de las relaciones establecidas. Una vez analizada la consistencia interna entre analistas, se aplica el sistema de categorías elaborado al análisis de tests sociométricos aplicados a alumnado de educación primaria y escolarizado en aulas multiculturales a fin de estudiar el nivel y naturaleza de la integración de diferentes minorías étnicas. El resto del artículo se destina a presentar los resultados obtenidos del análisis de la integración de la muestra de 95 alumnos y alumnas pertenecientes a distintas minorías étnicas y a describir las conclusiones y límites del mismo.

ABSTRACT

That contribution shows a categorisation system in order to analyse the psicosocial setup of people inside a group. It implies a kind of information analysis more qualitative and dynamic than the traditional and well known sociometric indexes. The categories present not only the quantity of recieved choices, rejections and perceptions, but the quality of established relations. Studied the intern consistency by different analistsds, the elaborated categorical system can by applied to the study of sociometric tests to the primary students in multicultural classroom, in order to study the level and nature of integration among different ethnic minorities. A least, that contribution ofers the outcomes from the analysis of the integration of a sample

of 95 students (girls and boys) from several ethnic minorities and describes their limits and conclusions.

I. INTRODUCCIÓN

La finalidad de este artículo es presentar un sistema de categorías para el análisis de la situación psicosocial de las personas dentro del sistema de relaciones de un grupo. En los estudios psicométricos normalmente se utilizan como indicadores del nivel de integración en un grupo algunos índices sociométricos en los que se cuantifica de alguna manera la posición de los sujetos; tomándose como referencia el recuento de las respuestas de un grupo a preguntas sobre sus preferencias, rechazos y respectivas percepciones a la hora de elegir compañeros y compañeras para realizar una actividad (Morales, 1975; Arruga, 1979; Bastin, 1973). La ventaja indudable de estos índices es su objetividad en la medida del nivel de integración, pero se reconoce que la información proporcionada por ellos es muy pobre respecto a la que se puede obtener a través de la interpretación más dinámica que ofrece la representación gráfica (sociograma) del sistema de relaciones del grupo en general y de los sujetos en particular.

Esta deficiencia de los tradicionales índices sociométricos nos animó a elaborar un sistema de categorías, de naturaleza más cualitativa que éstos, lo suficientemente exhaustiva y excluyente como para reflejar la posición de los sujetos dentro del sistema de relaciones del grupo. En este sistema de categorías se contempla no sólo el nivel de elecciones y rechazos que reciben los sujetos, sino también la calidad de las relaciones que establecen con el grupo y de éste con ellos.

El sistema de categorías elaborado se ha aplicado al análisis de los tests sociométricos aplicados a una muestra de alumnado de Educación Primaria escolarizados en aulas multiculturales a fin de estudiar el nivel de integración de diferentes minorías étnicas. El sistema de categorías elaborado nos permite analizar no sólo el nivel de integración de las minorías dentro de la dinámica de relaciones en sus grupos-clase, sino también la naturaleza de esta integración.

Se trata de un sistema de análisis de resultados sociométricos que permite responder a preguntas tales como ¿en qué medida las minorías se muestran abiertas al grupo-clase o prefieren relacionarse sólo con el alumnado de su propia minoría?, ¿se tratan de alumnos/as aislados/as o ignorados/as en el grupo?, ¿qué actitud tiene la clase hacia las minorías? ¿se muestra receptiva, rechaza o ignora a estos colectivos?, ¿en qué medida la imagen que tiene el alumnado de las minorías sobre su posición en el grupo corresponde con la que el grupo les devuelve? Las respuestas a estas preguntas y a otras que se concretarán a lo largo del trabajo ofrecerán una visión del nivel y naturaleza de la integración de los/as alumnos/as de las minorías dentro de la dinámica de relaciones que se establecen en sus respectivos grupos-clase.

El sistema de categorías elaborado representa un paso más allá en la interpretación de los resultados del test sociométrico respecto a los trabajos realizados por Davey (1982), Denscombe y otros (1986), Thirkell y Worrall (1989) y Díaz Aguado y Baraja

(1993) con minorías étnicas y los de Asher Hymel (1981), Conoley y Conoley (1983), Connolly (1983) en otros contextos, en los cuales la variable integración social se mide exclusivamente mediante el índice «estatus sociométrico» (cantidad de elecciones y rechazos recibidos).

2. DESCRIPCIÓN DEL TEST SOCIOMÉTRICO

El test sociométrico elaborado incluye las siguientes preguntas:

- Preferencia positiva: *De tus compañeros o compañeras de clase ¿con quién o quiénes te gustaría sentarte en autobús cuando hacéis una salida?* Con esta pregunta pretendemos conocer las preferencias de relaciones que manifiesta el alumnado de las minorías respecto a los de la mayoría y a la inversa, así como sus situaciones en el sistema de relaciones afectivas que se establecen en el grupo.
- Percepción de preferencia: *¿Quién o quiénes de tus compañeros y compañeras crees que te han elegido a ti para sentarse contigo en el autobús cuando hacéis una salida?* Los resultados de esta pregunta permiten conocer en qué medida la imagen que tienen los sujetos sobre su posición en el grupo, se corresponde con la que tienen en realidad, es decir, con la que el grupo les devuelve. Permitirá hacer un diagnóstico del mayor o menor realismo perceptivo del alumnado respecto a las relaciones afectivas positivas que mantiene en su clase.
- Rechazo: *¿Con quién o quiénes no te gustaría sentarte en el autobús cuando hacéis una salida?* Con esta pregunta pretendemos conocer las relaciones de antipatía que se establecen en el grupo y la posición de las minorías en ellas.
- Percepción de rechazos: *¿A quién o quiénes crees que no le gustaría sentarse contigo en el autobús cuando hacéis una salida?* Los resultados de esta pregunta permiten diagnosticar si la minoría se siente o no rechazada por la mayoría, y su realismo perceptivo respecto a las relaciones de antipatía que mantiene con el grupo.

Como puede observarse se trata de preguntas afectivas sobre preferencias deseadas y no sobre hechos. Hemos considerado que para los objetivos que perseguimos era más interesante el análisis de las relaciones de simpatía y antipatía latentes en el grupo, expresadas por los deseos de preferencias y rechazos que manifiesta el alumnado, que las establecidas entre ellos, que muchas veces son impuestas por las mismas condiciones físicas del contexto.

Hemos optado por elecciones ilimitadas en todas las preguntas (el alumnado puede elegir en cada una de ellas a todos los compañeros/as que quieran) para obtener una mayor espontaneidad en sus respuestas y determinar la real expansividad del alumno y la alumna hacia el grupo, tanto la de signo positivo como la de negativo. No obstante, hemos utilizado en la pregunta una actividad —*sentarse en el autobús*— que no favorece que se hagan muchas elecciones, procurando así una mayor discriminación entre una verdadera expansividad y una expansividad animada por la actividad que se propone en la pregunta.

El cuestionario sociométrico se imprimió en lengua castellana y catalana a fin de adaptarlo a la lengua de aprendizaje, quedando a discreción del alumno el utilizar una u otra versión.

3. APLICACIÓN DEL TEST SOCIOMÉTRICO

La aplicación de los tests sociométricos fue realizada por personal colaborador de la investigación. Estas personas fueron preparadas especialmente para ello y prácticamente en todas las aplicaciones estaban presentes dos de ellas a fin de controlar y optimizar las condiciones en que se realizaban la aplicación, atendiendo a las necesidades especiales del alumnado, especialmente el de las minorías étnicas (por ejemplo, no saber el castellano o el catalán).

Las personas colaboradoras debían realizar una memoria que recogiera todas las circunstancias e incidencias que ocurrieran durante la aplicación. Era imprescindible consultar dicha memoria para dar o no credibilidad a los resultados, o para interpretarlos con ciertas cautelas. A pesar de que se procuró controlar las condiciones de aplicación, en algunos colegios fue totalmente imposible conseguir unas condiciones mínimamente aceptables por falta de colaboración e indisciplina del alumnado de la clase en cuestión, por lo que tuvieron que ser desechados de la muestra definitiva.

4. ELABORACIÓN DEL SISTEMA DE CATEGORÍAS SOCIOMÉTRICAS

El proceso hasta llegar al sistema de categorías definitivo que presentamos más abajo, es el propio de estos casos. Al principio se elaboró un bosquejo de sistema, posteriormente se analizaron las respuestas del alumnado de las minorías de tres clases lo que permitió dejarlo muy acabado, de tal manera que en el transcurso del análisis no hubo necesidad de añadir ninguna categoría más para recoger algún aspecto significativo que afectara a un número importante del alumnado; aunque sí algún aspecto que, siendo de cierto interés señalarlo, sólo afecta a algunos casos particulares, como se verá en los comentarios de los resultados.

El sistema elaborado lo conforman las siguientes 9 dimensiones sociométricas y sus categorías subyacentes.

- A. *Nivel de expansividad positiva de las minorías hacia la clase*: recoge el nivel de apertura de las minorías hacia el grupo y hacia quién se dirige esta expansividad. Se realiza a partir del análisis de las elecciones efectuadas por las minorías. Se trata de una dimensión que nos indicará el interés de las minorías por relacionarse con el grupo.
 - A.1. Baja expansividad: no hace ninguna elección o muy pocas comparadas con las que hacen el resto de sus compañeros. A su vez, presenta tres subniveles:
 - A.1.1. No elige a nadie o elige a alguien marginado.
 - A.1.2. Elige a muy pocos pero indistintamente dirigidas, es decir, no concentra sus preferencias en compañeros/as de la misma minoría.

- A.1.3. Bajo nivel de elecciones y, además, dirigidas tan sólo a compañeros o compañeras de su misma minoría.
 - A.2. Nivel normal del expansividad: presenta un número de elecciones semejante al de sus compañeros/as de clase y se dirigen de forma indistinta, es decir, eligen tanto a compañeros/as de su misma minoría como a los de la mayoría.
 - A.3. Alto nivel de expansividad: sujetos que eligen a muchos compañeros o compañeras, destacando por encima de lo normal de elecciones que se hacen en su clase.
- B. *Nivel de receptividad del grupo-clase hacia las minorías:* recoge el comportamiento que expresa la clase hacia la minoría a partir del análisis de las elecciones efectuadas por el grupo.
- B.1. Bajo nivel de elecciones: no recibe elecciones o muy pocas (1 ó 2) por parte del grupo. Se trata de una persona situada en el círculo exterior del sociograma, sea en la figura de ignorado sea en la de olvidado. A la vez, presenta tres niveles:
 - B.1.1. No recibe ninguna elección o sólo una de algún marginado.
 - B.1.2. Recibe muy pocas elecciones pero no son especialmente de su minoría o de algún marginado.
 - B.1.3. Recibe muy pocas elecciones y son fundamentalmente de su minoría o de alguien perteneciente a alguna minoría.
 - B.2. Nivel de elecciones recibidas normal dentro de su grupo, e indistintamente recibidas, es decir, le eligen tanto alumnos/as de su propia minoría como de la mayoría.
 - B.3. Nivel alto de elecciones recibidas: lo eligen muchos compañeros/as, destacándose de la mayoría de sus compañeros/as por el mayor número de elecciones recibidas.
- C. *Nivel de expansividad negativa de las minorías:* recoge el grado de antipatía que expresan las minorías hacia la clase mediante el análisis de los rechazos que emiten.
- C.1. No rechaza a nadie o a uno insignificante: recoge principalmente los casos donde no se ha emitido rechazo o si se ha hecho es a un compañero o compañera que por su posición en el grupo no parece significativo tal rechazo.
 - C.2. Igual que otros compañeros/as: se trata de alumnado que emite un número de rechazos similar al que hace la mayoría de la clase, sea que los emita a toda la clase, sea que los emita al sexo contrario; un caso muy frecuente es el elevado número de rechazos que emiten las niñas hacia los niños independientemente de la minoría.
 - C.3. Elevado nivel de emisión de rechazos: es alumnado que emite un número de rechazos superior a los que hacen la mayoría de sus compañeros/as.

- C.4. Rechaza sólo a los de su propia minoría: categoría que se añadió en el transcurso del análisis creyéndose que podrían encontrarse algunos en esta situación, pero sólo recoge el caso particular que planteó su necesidad, puesto que esta situación no se da en el alumnado analizado.
- D. *Nivel de rechazo hacia las minorías por parte del grupo*: recoge el grado en que el grupo rechaza a las minorías a partir del análisis de los rechazos emitidos por sus miembros. Muestra el nivel de antipatía que expresa el grupo hacia las minorías.
- D.1. No recibe rechazos o uno poco significativo: incluye aquél que no recibe ningún rechazo o si recibe alguno es de alguien que por su posición en el grupo no es un rechazo significativo (por ejemplo, lo recibe de un alumno/a que se ha dedicado a rechazar a muchos).
- D.2. Recibe un nivel de rechazo igual que sus compañeros/as: recibe un número de rechazos similar al que reciben sus compañeros/as, ya sean del grupo total ya sea de su mismo sexo; es frecuente que los niños tengan un mayor número de rechazos que las niñas porque éstas suelen rechazarlos bastante.
- D.3. Recibe un alto nivel de rechazo: tiene un número de rechazos netamente superior al que reciben sus compañeros o compañeras.
- D.4. Sólo recibe rechazos de compañeros/as de su propia minoría: categoría que se añadió en el transcurso del análisis creyéndose que podrían encontrarse algunos en esta situación, pero sólo recoge algún caso muy particular.
- E. *Nivel de integración de las minorías en el grupo*: recoge la apreciación global de la situación de las minorías dentro del grupo a partir del análisis de sus posiciones dentro del sociograma grupal (representación gráfica de los resultados del test sociométrico) y de las relaciones que establecen con sus compañeros o compañeras.
- E.1. Sin integración.
- E.1.1. Aislados: alumnos/as que no realizan ni reciben ninguna elección.
- E.1.2. Sin integración relacionándose tan sólo con algún marginado o poco integrado como él, que no sea de su misma minoría.
- E.1.3. Sin integración y sólo relacionado con alguno de su minoría o con deseos de relacionarse con ella: sólo se abre hacia los compañeros/as de su misma minoría.
- E.1.4. Sin integración pero abierto al grupo: son alumnos o alumnas ignorados por la clase en cuanto que ellos emiten elecciones a compañeros/as relativamente bien situados en el grupo e independientemente de su minoría, pero no reciben ninguna elección o tan sólo una muy poco significativa en el contexto de sus relaciones (puede ser de alguien que ha elegido prácticamente a toda la clase o todos los de su género como a veces suele suceder).
- E.2. Bajo nivel de integración: alumno/a situado en los círculos externos del sociograma, pero con alguna relación; no reciben muchas elecciones pero sí algunas que pueden ser o no recíprocas.

- E.3. Buen nivel de integración:
 - E.3.1. Bien relacionado: alumnado situado en los círculos medios y centrales del sociograma con un buen sistema de relaciones.
 - E.3.2. Populares: alumnado situado en el círculo central y con un nivel de elecciones netamente superior a las recibidas por la mayoría de sus compañeros/as.

- F. *Percepción de preferencias de las minorías*: recoge las expectativas de elección positiva de las minorías a partir de analizar las preguntas de percepción de preferencias positivas. El interés de esta dimensión está en indicar la imagen que tienen las minorías de sus relaciones afectivas con el grupo.
 - F.1. Bajo nivel de expectativas: los alumnos/as dejan en blanco la respuesta o indican muy pocos compañeros comparados con los que indican el resto del grupo.
 - F.1.1. Expectativas prácticamente nulas: recoge los casos en que los alumnos/as no contestan a la pregunta o indican expresamente que no saben quién les ha elegido, o si lo hacen tan sólo indican uno que no parece ser significativo dentro de su sistema de relaciones.
 - F.1.2. Bajo nivel de expectativas y además referidas sólo a los de su minoría: el alumno/a cree que sólo lo elegirán compañeros/as que pertenecen a su misma minoría.
 - F.1.3. Bajo nivel de expectativas pero indistintamente distribuidas: creen que les eligirá uno o dos compañeros o compañeras (cuando en términos generales la mayoría del grupo indican más) pero no necesariamente perteneciente a su propia minoría.
 - F.2. Nivel de expectativas normal: el alumno/a cree que va a tener un número de elecciones similar al que expresan sus compañeros/as del grupo.
 - F.3. Nivel alto de expectativas: cree que será elegido por un número mayor de los que expresan sus compañeros/as de clase.

- G. *Percepción de rechazos de las minorías*: se recoge las expectativas de rechazos de las minorías a partir de analizar la pregunta sobre percepción de rechazos. Proporciona la percepción que tiene el alumnado sobre las relaciones de antipatía que mantiene con sus compañeros/as. En qué grado se considera más o menos rechazado.
 - G.1. Bajo nivel de percepción de rechazos: son aquellos casos donde dejan sin respuesta la pregunta o tan sólo indican que lo rechazará uno que no parece significativo en su sistema de relaciones.
 - G.2. Nivel de percepción de rechazo normal dentro de su grupo: creen que tendrán un número de rechazos semejante al que expresa la mayoría de sus compañeros/as, sea del grupo general, sea de su mismo grupo sexual.
 - G.3. Alto nivel de percepción de rechazo: cree que va a ser muy rechazado.

- H. *Realismo perceptivo de preferencias*: recoge en que medida coincide la imagen que el alumno/a tiene de su sistema de relaciones positivas con la que el grupo expresa sobre ellas. Se diagnostica a partir de contrastar las elecciones de preferencias positivas que recibe el alumno/a con las que cree que va a recibir, según expresa en la pregunta de percepción correspondiente.
- H.1. Sí: existe una gran coincidencia entre las elecciones que el alumno/a de la minoría espera recibir de sus compañeros y las que realmente recibe, tanto por lo que se refiere al número como a la dirección de las mismas.
- H.2. No: hay muy poca coincidencia entre las elecciones que el alumno/a espera recibir de la clase y las que realmente recibe, tanto por lo que se refiere al número de elecciones como a la dirección de las mismas.
- H.3. Ambiguo: incluye los casos donde no se puede apreciar con claridad si el alumno/a tiene o no un realismo perceptivo, bien porque deja algunas preguntas en blanco, bien porque en parte coincide las elecciones que espera recibir y las que recibe y en parte no coincide.
- I. *Realismo perceptivo de rechazos*: recoge en qué medida coincide la imagen que el alumno o alumna tiene de su sistema de relaciones de antipatía con la que el grupo le devuelve. Se diagnostica a partir de contrastar los rechazos que recibe el alumno/a con los que él cree que va a recibir, según expresa en la pregunta de percepción correspondiente.
- I.1. Sí: existe un gran coincidencia entre los rechazos que el alumno/a de la minoría espera recibir y los que realmente recibe, tanto por lo que se refiere al número como a la dirección de los mismos.
- I.2. No: hay muy poca coincidencia entre los rechazos que el alumno o alumna espera recibir de la clase y los que realmente recibe, tanto por lo que se refiere al número de rechazos como a la dirección de los mismos.
- I.3. Ambiguo: incluye los casos donde no se puede apreciar con claridad si el alumno/a tiene o no un realismo perceptivo sobre los rechazos, bien porque deja algunas preguntas en blanco, o bien porque en parte coinciden los rechazos que espera y los que recibe y en parte no.

5. ESTUDIO DE LA CONSISTENCIA ENTRE ANALISTAS

Uno de los problemas de fiabilidad más graves que se presenta en la metodología del Análisis de Contenido es la consistencia entre los analistas en el momento de asignar las unidades de análisis a una u otra categoría. El no asegurar este aspecto puede invalidar los resultados hallados. En nuestro caso, el análisis sociométrico a realizar era bastante complejo puesto que, aún siendo exhaustivas las definiciones, la mayoría de las categorías debían contemplarse en el contexto más amplio del grupo. Es decir, la situación que caracterizaba al alumnado de la minoría para asignarle a uno u otro nivel de las categorías, dependía del sistema de relaciones del grupo y del comportamiento de la mayoría de sus miembros.

Condiciones que se tuvieron en consideración a fin de asegurar la consistencia interna entre los analistas fueron:

- a) Los analistas fueron las mismas personas que elaboraron el sistema de categorías lo que garantiza que su interpretación sea igual, además de la experiencia de analizar conjuntamente a varios alumnos/as (11 alumnos/as pertenecientes a tres clases).
- b) Cuando se estableció el sistema de categorías definitivo se realizó un estudio sobre el índice de acuerdo entre los analistas. Se eligieron cuatro colegios y al azar dentro de ellos a dos alumnos/as de las minorías que fueron analizados por cada una de ellas por separado. Posteriormente se calculó el índice de acuerdo entre ambas categorizaciones. Se obtuvo un índice de acuerdo del 96%.

Consideramos que está asegurada la fiabilidad del análisis por lo que se refiere a la consistencia entre analistas, a tenor de las precauciones adoptadas en la elaboración del sistema de categorías y del valor obtenido en el índice de acuerdo entre los analistas que realizarán el estudio sociométrico.

6. DESCRIPCIÓN DE LA MUESTRA

El total de alumnos/as sobre los que se realiza el estudio de su nivel de integración en el grupo clase es de 95. De la muestra inicial se eliminaron:

- Las aulas en donde el alumnado de estas minorías iguala o supera al de la cultura de acogida. Nos limitamos a las aulas donde las minorías eran realmente una minoría (igual o menor al 20% de la clase).
- Alumnado del 1º y 2º curso puesto que sus niveles de dominio en la lectura y escritura no garantizaban una aplicación colectiva eficaz del test sociométrico.
- Aquellas clases donde las condiciones de aplicación, a pesar de las precauciones adoptadas, no garantizan ni la fiabilidad ni la validez de sus resultados.
- Los alumnos/as que no pudieron hacer el test sociométrico por estar ausentes el día en que se aplicó o porque no se pudo superar sus problemas lingüísticos (en algunos niños mediante aplicación individual sí se consiguió).

En consecuencia, los resultados del nivel de integración en el aula se realizan con una muestra de 95 alumnos y alumnas de minorías étnicas que dentro de sus aulas constituyen grupos minoritarios, escolarizados en los cursos de 3º a 6º en escuelas ubicadas en medios socioeconómicos desfavorecidos. En total han participado 13 escuelas. Respecto a las minorías presentes en el estudio, la magrebí es la que tiene una real representación en nuestra muestra (63), y luego la filipina (14), pero las otras minorías (guineana, peruana, dominicana y paquistaní) están representadas por 2 o tres alumnos o alumnas por ello las hemos considerado para el análisis como un único bloque (13).

Otras variables que se han considerado para caracterizar la muestra, porque pueden proporcionar patrones de integración en el aula diferenciales, son los siguientes:

- Sexo: están equilibradas las representaciones de los dos géneros: el 56 % son alumnos y el 44% alumnas.
- Dominio de la lengua castellana: la mayoría del alumnado tienen un cierto dominio de la lengua castellana, el 59% la domina bien, el 27% regular, y sólo un 13 % tiene dificultades de comprensión y expresión.
- Años de ingreso en el colegio: hay un predominio de alumnado que llevan más de tres años en el colegio (un 49%), fácilmente explicable dado los cursos escolares que caracterizan a la muestra. El 25% es alumnado que han ingresado en el centro el mismo curso que se realiza la investigación (1993-94), y el 22% llevan dos o tres años.

7. INTEGRACIÓN EN EL AULA DE LAS MINORÍAS ÉTNICAS. ANÁLISIS DE RESULTADOS

Como paso previo e imprescindible para aplicar el sistema de categorías elaborado se confeccionaron las sociomatrices y sociogramas de cada uno de los grupo-clase. A continuación se exponen los resultados obtenidos en cada una de las nueve dimensiones sociométricas y sus categorías subyacentes.

A. Nivel de expansividad positiva de las minorías. Emisiones de atracción hacia el grupo

La tabla 1 recoge el nivel de expansividad hacia el grupo o emisiones de atracción del alumnado de las minorías, agrupando subcategorías.

Tabla 1

EMISIONES DE ATRACCIÓN DE LAS MINORÍAS HACIA EL GRUPO-CLASE

MINORÍAS	BAJA EXPANSIVIDAD				MEDIA o ALTA EXPANSIVIDAD			TOTAL
	(1)	(2)	(3)	Subtotal	(4)	(5)	Subtotal	
Magrebíes	11%	23%	66%	41%	89%	11%	61%	63
Filipinos	—	29%	71%	50%	86%	14%	50%	14
Otras minorías	17%	33%	50%	46%	71%	29%	54%	13
Subtotal	4%	11%	28%		49%	8%		
Total	43%				57%			90

(1) No elige a nadie o algún marginado/a.

(2) Bajo nivel de elecciones pero indistintamente dirigidas.

(3) Elige a pocos/as y sólo de su minoría.

(4) Nivel de elecciones normal, semejante a las de su grupo clase e indistintamente dirigidas.

(5) Elevado nivel de elecciones e indistintamente dirigidas.

Es significativo el dato de que un 43% del alumnado de las minorías tiene un nivel bajo de expansividad. Entre los que presentan bajo nivel de expansividad cabe resaltar que el 64% de ellos dirigen sus elecciones sólo a niños de su propia minoría sin abrirse a otros miembros del grupo (suponen un 28% del total). A este patrón de comportamiento responden las distintas minorías.

B. Nivel de receptividad del grupo-clase hacia las minorías. Elecciones recibidas

En la tabla 2 podemos ver que algo más de la mitad del alumnado de la minoría recibe un bajo nivel de elecciones por parte del grupo (54%). Sin embargo en el nivel de acogimiento de estos niños/as por parte del grupo se observan ciertos aspectos diferenciales. Así, entre los niños y niñas magrebíes hay un 58% que tienen un bajo nivel de recepción de elecciones, frente a sólo un 36% en el caso del alumnado filipino. En estos últimos niños, observamos también, no sólo que el 64% de ellos tienen un buen nivel de receptividad por parte del grupo, sino que además hay 1/3 de ellos que son populares.

Tabla 2

RECEPCIÓN DE ELECCIONES DE LAS MINORÍAS, POR PARTE DEL GRUPO-CLASE

MINORÍAS	NIVEL DE RECEPCIÓN MUY BAJO				NIVEL DE RECEPCIÓN MUY ALTO			TOTAL
	(1)	(2)	(3)	Subtotal	(4)	(5)	Subtotal	
Magrebíes	61%	3%	31%	58%	82%	18%	42%	67
Filipinos	60%	—	40%	36%	67%	33%	64%	14
Otras minorías	57%	—	43%	50%	86%	14%	50%	14
Subtotal	33%	3%	19%		37%	10%		
Total	54%				46%			95

(1) No recibe elección o sólo de algún marginado/a.

(2) Recibe pocas elecciones, pero no son de marginados/as ni de su minoría.

(3) Recibe muy pocas elecciones y sólo de su minoría o de alguien de la minoría, sean o no marginados/as.

(4) Nivel medio de recepciones, pero de procedencia indistinta.

(5) Nivel alto de elecciones recibidas.

Por último, cabe resaltar que 1/3 del grupo total del alumnado no recibe ninguna elección o sólo de algún niño o niña marginado, y esto se produce, sobre todo, en el caso de los magrebíes, en donde hay un 36%.

C. Nivel de expansividad negativa de las minorías respecto al grupo-clase. Emisiones de rechazo

De los resultados del análisis de la tabla 3 se desprende que la mayoría del alumnado de estas minorías o no emite rechazos a otros niños o niñas (35%) o bien lo hacen

igual que otros (47%). No se da un rechazo entre el alumnado de la misma minoría y tan sólo hay un niño que ha dirigido su rechazo de forma exclusiva a alguien de su propia minoría.

Tabla 3
EMISIONES DE RECHAZO DE LAS MINORÍAS HACIA EL GRUPO-CLASE

MINORÍAS	No rechazan a nadie o uno insignificante	Igual que otros compañeros/as	Rechazan más que lo normal de su grupo	Rechazan sólo a alguno/a de su minoría	TOTAL
Magrebíes	37%	51%	12%	—	63
Filipinos	36%	43%	21%	—	14
Otras minorías	31%	31%	31%	7%	13
Total	35%	47%	17%	1%	90

Si nos fijamos en el alumnado que emite un nivel de rechazos superior a lo normal de su clase, vemos que comparativamente los del grupo magrebí rechaza menos que los niños filipinos o de las otras minorías estudiadas. Esta tendencia, sin embargo, hay que interpretarla con mucha cautela, dado el reducido número de niños y niñas que hay de otras minorías.

D. Nivel de rechazo hacia las minorías, por parte del grupo-clase. Recepción de rechazos

En la tabla 4 que se presenta a continuación, se puede apreciar cómo los niños y niñas de estas minorías son alumnos y alumnas que no concentran los rechazos de la clase. En su mayoría (69%) no reciben rechazos o bien los reciben igual que sus compañeros/as de grupo.

Tabla 4
RECEPCIÓN DE RECHAZOS DE LAS MINORÍAS

MINORÍAS	No reciben rechazo a uno insignificante	Igual de rechazos que los de su grupo	Reciben alto nivel de rechazos	Reciben sólo de su minoría	TOTAL
Magrebíes	33%	34%	28%	5%	67
Filipinos	50%	36%	14%	—	14
Otras minorías	29%	29%	42%	—	14
Total	35%	34%	28%	3%	95

Cabe, sin embargo, destacar que, entre el alumnado que es rechazado por el grupo (28%), el grupo filipino es el que recibe menos rechazos por parte de sus compañeros/as.

E. Nivel de integración del alumnado de las minorías en su grupo

El análisis de los resultados de las categorías de esta dimensión confirman las apreciaciones comentadas hasta ahora en las dimensiones anteriores, sobre todo en lo referente a las dos primeras: la expansividad del alumno/a hacia la clase y la actitud de la clase hacia ellos.

Tabla 5
NIVEL DE INTEGRACIÓN DEL ALUMNADO DE LAS MINORÍAS

MINORÍAS	SIN INTEGRACIÓN		BAJO NIVEL DE INTEGRACIÓN	BUEN NIVEL DE INTEGRACIÓN		TOTAL
	Aislados	Sin integración con deseos de integración		Bien relacionado	Populares	
Magrebíes	6%	47%	14%	25%	8%	64
Filipinos	—	23%	23%	31%	23%	13
Otras minorías	14%	14%	21%	30%	21%	14
Total	45%	17%	38%	91		

Como puede observarse en la tabla 5, el 62% de los alumnos/as de estas minorías se encuentran principalmente en niveles de integración muy bajos, y de ellos un 45% sin integración. Ahora bien, esto debe matizarse por minorías. Así, en el caso de los niños y niñas magrebíes hay un 53% que no están integrados, mientras que en el alumnado filipino es sólo del 23%. Esa falta de integración se produce no porque ellos tengan un comportamiento propio de niños/as aislados/as (no eligen ni son elegidos/as) sino que se trata sobre todo de niños y niñas ignorados por el grupo, es decir, expresan su deseo de entrar en contacto con el grupo mediante sus elecciones, pero no reciben elecciones de sus compañeros/as.

Parece que el alumnado filipino y de otras minorías tienen un mayor nivel de integración, pues algo más de la mitad de ellos presentan una buena integración.

F. Percepción de preferencias. Expectativas o creencias de su situación de aceptación dentro del grupo

En la tabla 6 se puede observar que el alumnado de las minorías cree que va a ser poco elegidos/as por sus compañeros/as (66%) y las elecciones que perciben pueden

venir tanto de su minoría como de otros niños/as. Se aparta de esta apreciación el caso de la minoría filipina que, coincidiendo con su mayor nivel de integración comentado anteriormente, creen que van a tener un nivel de elección igual que sus compañeros y compañeras e incluso alguno y alguna cree que será muy elegido/a (46%).

Tabla 6
PERCEPCIÓN DE PREFERENCIAS DEL ALUMNADO DE LAS MINORÍAS

MINORÍAS	BAJA			MEDIA	ALTA	TOTAL
	(1)	(2)	(3)	(4)	(5)	
Magrebíes	25%	18%	24%	27%	6%	63
Filipinos	8%	23%	23%	38%	8%	14
Otras minorías	38%	8%	31%	23%	—	13
Subtotal	25%	16%	25%	28%	6%	
Total		66%		28%	6%	89

- (1) No saben quien puede elegirle o alguno/a insignificante.
- (2) Cree que va a ser sólo elegido/a por los de su minoría.
- (3) Cree que va a ser elegido/a por alguno que no es de su minoría.
- (4) Cree que va a ser elegido/a de forma indistinta, sin diferencias del grupo.
- (5) Cree que va a ser muy elegido/a.

G. Percepción de rechazos. Creencias de su situación de rechazo dentro del grupo

Llama la atención el hecho de que el grupo mayoritario presenta un nivel de rechazo mayor hacia los niños y niñas de estas minorías que lo que ellos piensan. Así, podemos ver en la tabla 7 que mientras el 8% del alumnado de las minorías creen que van a ser muy rechazados/as, el 28% de ellos han recibido un alto nivel de rechazo, tal como se comentó al tratar la recepción de rechazo de estas minorías.

El alumnado de las minorías no tienen conciencia de ser rechazados —o no saben, o no creen que van a ser rechazados igual que la mayoría—. A destacar es el hecho de que mientras el grupo filipino cree en su mayoría que van a tener un nivel de rechazo como el resto de sus compañeros (64%), los magrebíes y las otras minorías dicen no saberlo o no creer ser rechazados o por alguien insignificante. Ello puede ser indicativo de que estos últimos tengan un mayor desconocimiento de su situación en el grupo o bien que tengan una percepción de rechazo menor.

Tabla 7
PERCEPCIÓN DE RECHAZO DE LOS NIÑOS DE LAS MINORÍAS

MINORÍAS	Percepción de rechazo			TOTAL
	BAJO	MEDIO	ALTO	
Magrebíes	52%	42%	6%	62
Filipinos	28%	64%	7%	14
Otras minorías	54%	31%	15%	13
Total	48%	44%	8%	89

(1) No lo sabe. No cree ser rechazado/a o tan sólo por alguno/a insignificante.

(2) Cree que tendrá un nivel de rechazo semejante a la mayoría del grupo.

(3) Cree que va a ser muy rechazado/a.

I. Realismo perceptivo positivo

En líneas generales podemos decir que la mitad de este alumnado tienen una percepción de su situación de atracción dentro del grupo que sí se corresponde con sus expectativas o creencias. La imagen que tienen de su situación en el grupo se corresponde con la que el grupo les devuelve, es realista. Sólo en el 27% de ellos no se produce una coincidencia.

Tabla 8
REALISMO PERCEPTIVO POSITIVO

MINORÍAS	Percepción de rechazo			TOTAL
	SI	NO	Ambiguo	
Magrebíes	46%	21%	25%	63
Filipinos	43%	36%	21%	14
Otras minorías	77%	8%	15%	13
Total	50%	27%	23%	90

J. Realismo perceptivo negativo

La distribución de frecuencias entre las distintas categorías de esta dimensión, como puede apreciarse en la tabla 9, nos muestra que la imagen que el alumnado tiene de su posición en el grupo no está clara, pues se reparten por igual.

Tabla 9
REALISMO PERCEPTIVO NEGATIVO

MINORÍAS	Percepción de rechazo			TOTAL
	SÍ	NO	Ambiguo	
Magrebíes	30%	29%	41%	63
Filipinos	42%	29%	29%	14
Otras minorías	46%	46%	8%	13
Total	34%	31%	34%	90

8. CONCLUSIONES GENERALES

Los niños y las niñas de estas minorías, en general, emiten pocas elecciones afectivas de atracción, y muchas de ellas las dirigen hacia niños/as de su propia minoría; en su mayoría reciben un reducido número de elecciones por parte del grupo-clase (sobre todo los magrebíes, mientras que los filipinos cuentan con mayor grado de aceptación); es un alumnado poco dado a emitir rechazos (tendencia algo más marcada en el caso de los magrebíes) y el grupo tampoco le rechaza (sobre todo el grupo filipino). Cabe resaltar, sin embargo, algunos aspectos diferenciales.

El alumnado magrebí es el que muestra menor grado de integración. La mitad de él no está integrado en el grupo, y se encuentra básicamente ignorado por sus compañeros y compañeras de la mayoría, cuando éstos emiten preferencias afectivas de atracción. Este alumnado muestra deseo de estar con los de la mayoría, pero éstos no le elige.

Son niños y niñas bastante conscientes de su situación de relación afectiva, creen que van a ser poco elegidos, como realmente es. Sus expectativas se orientan tanto hacia los de su propia minoría como hacia el alumnado perteneciente al grupo mayoritario, de ahí que su realismo perceptivo positivo no sea tan acertado. No forman entre ellos un grupo cerrado.

No tienen sentimiento de rechazo por parte del grupo, pero reciben un nivel de rechazo mayor que el que ellos creen, pero como ya se ha comentado, éstos no son importantes. En los casos, muy particulares, en donde aparece que reciben rechazo por parte de los de su misma minoría, se da por parte de alguna niña respecto a algún niño.

El alumnado filipino tiene buen nivel de integración. Dentro de las minorías es el que está más integrado. Según se desprende de las respuestas obtenidas, su nivel de apertura hacia el grupo es semejante al de sus compañeros y compañeras del grupo clase. Dirige sus preferencias de forma indistinta, aunque las reciprocidades o relaciones mútuas son más frecuentes entre los de su propia minoría (entre las niñas, sobre todo, forman a menudo tríadas y parejas). Este alumnado tiene un buen nivel de acogimiento por parte de sus compañeros y compañeras y no es rechazado, y cuando esto ocurre, es como el resto del grupo.

Los niños/as son conscientes de su situación. Creen que van a ser elegidos de forma indistinta, aunque más por los de su minoría, y que pueden ser rechazados como los demás. Tienen mayor realismo perceptivo, tanto positivo como negativo.

En cuanto al alumnado de las restantes minorías estudiadas, es difícil extraer alguna conclusión dado el número tan reducido de niños y niñas de cada minoría de los que disponemos datos sociométricos. En los aspectos generales señalados para todo el alumnado de las minorías, se encontrarían en la misma situación que los demás, y en los aspectos diferenciales su situación parece estar en una posición intermedia entre el alumnado magrebí y el filipino.

No han resultado significativas las asociaciones realizadas entre los niveles de integración del alumnado y algunas variables que consideramos de interés. En concreto, se ha estudiado mediante tablas de contingencia la significación estadística (chi cuadrado) de posibles relaciones entre niveles de integración en el aula y los años de estancia en el centro de los alumnos y alumnas, o su género de pertenencia o su dominio de la lengua instrumental (catalán o castellano).

A fin de salvar el problema de celdillas vacías en las tablas de contingencias, se ha realizado una agrupación de las categorías consideradas en cada ámbito. Así, para cada aspecto sociométrico sólo se distinguieron tres niveles de integración: a) alto; b) igual que la mayoría; y c) bajo. Respecto a la variable años de permanencia en el centro, las categorías contempladas fueron: a) el primer año en el centro; b) de dos a tres años; y c) más de tres años. En el caso del dominio de la lengua sólo se distinguieron dos categorías: a) dominio suficiente para comunicarse con los iguales; y b) sin dominio o con claras dificultades de expresión.

9. LÍMITES DEL ESTUDIO

Las conclusiones precedentes deben interpretarse a la luz de las limitaciones inherentes a la propia técnica sociométrica utilizada y a las características de las muestras de las diferentes minorías.

En general, los errores de validez y fiabilidad de los resultados de un test sociométrico pueden tener su origen fundamentalmente en las preguntas, más concretamente, en la actividad que incluye, en las condiciones de aplicación y en la interpretación de los resultados, sobre todo cuando éstos son de naturaleza cualitativa más que cuantitativa.

Piénsese que los resultados obtenidos se han inferido de las respuestas que el alumnado ha dado en un momento preciso a las preguntas del test sociométrico. Podemos considerar que éstas, por la actividad que implicaban (*sentarse en el autobús*), motivaban respuestas de naturaleza afectiva tal como han sido interpretadas. No obstante, el momento preciso en que fue aplicado el test sociométrico (no se pudo controlar, quedaba a discreción de la escuela), la situación de los sujetos durante la aplicación (el que pudieran contestar de forma independiente unos de otros, por ejemplo), los estados de ánimos individuales y colectivos del momento, la propia sinceridad de los alumnos/as en sus respuestas, son fuentes de posibles errores que afectan a la interpretación de los resultados obtenidos.

Preocupados por minimizar en lo posible estas fuentes de error hemos intentado controlar al máximo las condiciones de aplicación, de tal manera que fueron eliminados los tests sociométricos de los grupos donde no se cumplieron los requisitos que garantizaban la credibilidad de los resultados. Ello nos hizo desistir de muchos grupos que en cierta medida pueden haber sesgado la muestra. Los grupos eliminados se caracterizaron sobre todo por la falta de disciplina del alumnado, y podemos preguntarnos: ¿los resultados hubiesen sido los mismos si se incluyen estos grupos conflictivos? ¿tiene alguna relación la conflictividad en el aula con las relaciones afectivas que mantienen las minorías y la mayoría entre sí?

No hay duda que la validez de nuestras conclusiones quedarían reafirmadas si las hubiésemos triangulado con la propia percepción del profesorado o con las obtenidas con la aplicación complementaria de otra técnica sociométrica. No obstante, consideramos que se han tomado suficientes medidas para justificar la validez de esta aproximación que ofrecemos sobre el nivel de integración de las minorías en el grupo clase. Por último, debemos señalar que un límite importante a tomar en consideración es el reducido número de alumnado de algunas minorías. Si bien el número de los magrebíes avala la consistencia de resultados hallados para este grupo, en el caso de los filipinos los resultados deben interpretarse con mucha cautela, y los referentes a las otras minorías sólo deben considerarse a modo complementario de los resultados globales.

BIBLIOGRAFÍA

- Asher, S.R. y Hymel, S. (1981). Children's social competence in peer relations: sociometric and behavioral assessment. En J.D. Wine y M.D. Smye (Eds.). *Social competence*. New York: Guilford Press.
- Besalú, X. y Palaudarais, J.M. (1994). Una recerca sobre nens immigrants estrangers a l'escola. En Varios. *Sobre interculturalitat. Girona. Documents de treball de la tercera i quarta Escola d'Estiu sobre interculturalitat*. (pp. 207-210). Girona: Fundació SER. GI).
- Cabrera, F.; Espín, J.; Marín, M.A. y Rodríguez, M. (1995). Estudio del nivel de integración de minorías étnicas en el aula. En AIDIPE (Comp.) *Estudios de investigación educativa en intervención psicopedagógica*. (pp. 147-150). Valencia: Aidipe.
- Conoley, J.C. y Conoley, C.W. (1983). A Comparison of techniques to affect sociometric status: a small step toward primary prevention in the classroom? *Journal of School Psychology*, 21, 41-47.
- Connolly, J.A. (1983). A review of sociometric procedures in the assessment of social competencies in children. *Applied Research in mental Eetardation*, 4, 315-327.
- Davey, A.G. (1982). Ethnic identification, preference and sociometric choice. En Gajendra K. Verma y Christopher Bagley (Eds.). *Self-concept, achievement and multicultural education*. (pp. 60-69). London: Macmillan.
- Díaz Aguado, M.J. y Baraja, A. (1993). *Interacción educativa y desventaja sociocultural*. Madrid: CIDE.
- Díaz Aguado, J.M. (1996). *Escuela y tolerancia*. Madrid: Pirámide.

- Driver, Geoffrey (1982). Ethnicity and cultural competence: Aspects of interaction in multiracial classrooms. En Gajendra K. Verma y Christopher Bagley (Eds.). *Self-concept, achievement and multicultural education*. (pp. 70-79). London: Macmillan.
- Jackson, P. (1975). *La vida en las aulas*. Madrid: Marova. (Edición original 1968).
- Jordán, J. A. (1992). *L'educació multicultural*. Barcelona: CEAC.
- Juliano, D. (1993). *Educación intercultural. Escuela y minorías étnicas*. Madrid: Eudema.
- Thirkell, B. y Worrall, N. (1989). Differential ethnic bias in bengali and white children. *Educational Research*, 31, 181-188.
- Verma, G.K. y Bagley, C. (1982). *Self-Concept, Achievement and Multicultural Education*. London: Macmillan Press.

LA BÚSQUEDA DE FACTORES PROTECTORES DEL FRACASO ESCOLAR EN NIÑOS EN SITUACIÓN DE RIESGO MEDIANTE UN ESTUDIO DE CASOS

Judit Fullana Noell

Departament de Pedagogia. Universitat de Girona

RESUMEN

En este artículo presentamos uno de los enfoques de la investigación que creemos que puede contribuir a hallar conocimientos útiles para la prevención de las situaciones de fracaso escolar. Se trata del enfoque que en algunos estudios viene a denominarse estudio del riesgo y la «resiliencia», término inglés que se refiere a la capacidad de los individuos para resistir acontecimientos adversos en su experiencia vital, sin consecuencias negativas o perturbadoras a largo plazo para su desarrollo o su socialización. Se presenta una revisión de estudios que se han realizado bajo este enfoque, y se ilustra con una investigación que tenía el propósito de identificar factores protectores del riesgo de fracaso escolar en adolescentes en situación de riesgo. Mediante un enfoque cualitativo de estudio de casos, esta investigación nos ha permitido aproximarnos a algunos factores personales, escolares y sociales que parecen tener cierta relevancia como factores protectores del riesgo de fracaso escolar.

ABSTRACT

The aim of this article is to introduce one of the approaches of research that we consider may contribute to find useful knowledge for prevention of school failure. This approach is in some studies called the study of risk and resilience. The word resilience refers to the capacity of individuals to endure adverse events in their vital experience, without long-term negative or disrupting consequences in their development or socialisation. We present a review of studies that have been carried out with this approach. Then it will be exemplified with a research that intended to identify protective factors of the risk of school failure in at risk ado-

lescents. By means of a qualitative case study approach, this research has allowed us to approach some personal, educational and social factors that seem to have a certain relevance as protective factors of the risk of school failure.

I. INTRODUCCIÓN

En este artículo presentamos uno de los enfoques de la investigación que creemos que puede contribuir a hallar conocimientos útiles para la prevención de las situaciones de fracaso escolar. Se trata del enfoque que en algunos estudios viene a denominarse estudio del *riesgo* y «resiliencia», del término inglés *resilience*¹. Este término hace referencia a *la capacidad de los individuos para resistir acontecimientos adversos en su experiencia vital, sin consecuencias negativas o perturbadoras a largo plazo para su desarrollo o su socialización*. Aunque este enfoque aplicado al estudio de los problemas de aprendizaje y de rendimiento escolar goza de una tradición de dos décadas de investigación en Estados Unidos, supone una cierta novedad en España.

Se presenta una revisión de estudios que se han realizado bajo este enfoque, y se ilustra con una investigación que, siguiendo el enfoque antes descrito, tenía el propósito de identificar factores protectores del riesgo de fracaso escolar en adolescentes que, a pesar de haber vivido circunstancias familiares muy problemáticas, habían conseguido alcanzar un cierto nivel de éxito escolar. Mediante un enfoque cualitativo de estudio de casos, esta investigación nos ha permitido aproximarnos a algunos factores personales, escolares y sociales que parecen tener cierta relevancia como factores protectores del riesgo de fracaso escolar.

2. LOS CONCEPTOS DE RESILIENCE Y DE FACTOR PROTECTOR

Antes de referirnos a las investigaciones que se han realizado bajo este enfoque es necesario definir que entendemos por *resilience* y por factor protector, ya que son los conceptos básicos que sirven para definir la perspectiva desde la cual diseñamos nuestra investigación.

Algunos autores como Luthar i Zigler (1991)² y Werner i Smith (1982)³ definen el concepto de *resilience* como la ausencia de psicopatologías o conductas desadaptadas. Sin embargo, coincidimos con otros autores en considerar que el término que tratamos de conceptualizar va más allá de la definición anterior. No hay que perder de vista que el concepto de *resilience*, tal como lo definen la mayoría de los investigadores, tiene sen-

1 La traducción literal del término inglés «resilience» sería *resistencia* o *capacidad para adaptarse*. Sin embargo, no existe un término equivalente en español para describir con exactitud el mismo fenómeno al que se refiere el término inglés.

2 Citado por Miller, M. (1995). *Sources of Resilience Outcomes*. Paper presented at the International Convention of the Council for Exceptional Children, Indianapolis, April 5-9-1995. Documento del Eric Document Reproduction Service (EDRS). Núm.: ED381 949, p. 3.

3 Citado por Miller, M. (1995). Op. cit. p. 3.

tido en la medida en que nos estemos refiriendo a niños o personas que deben enfrentarse a situaciones adversas, a tensiones graves en la infancia o la adolescencia, en definitiva, que han sufrido o sufren aún situaciones de riesgo. Así, para Rutter (1990)⁴ el término *resilience* hace referencia a la *capacidad de los individuos para resistir acontecimientos adversos en su experiencia vital, sin consecuencias negativas o perturbadoras a largo plazo para su desarrollo o su socialización*. Para Manciaux (1994) un niño «resiliente» es el que en condiciones difíciles y desestabilizadoras es capaz de continuar su camino, se comporta de forma eficaz y llega a ser, de este modo, un adulto competente. Pare Beardslee (1989)⁵ *resilience* implica la buena adaptación inusual ante tensiones severas y para Berliner y Benard, (1995) este término se refiere al potencial de los jóvenes para desarrollarse en adultos saludables, productivos y competentes a pesar de las experiencias y tensiones severas y las adversidades que deben superar. Gordon, Ingersoll y Orr (1995) la conceptualizan como una habilidad para tener éxito, madurar y ganar competencia en un contexto de circunstancias adversas o de obstáculos. Por consiguiente hay que tener en cuenta que el término «*resilience*» *alcanza su máxima significación cuando se hace referencia a las respuestas de los individuos ante situaciones de riesgo, respuestas que incluyen aspectos conductuales, afectivos y cognitivos*.

Este concepto tiene, pues, dos dimensiones a considerar. Una, viene determinada por la *situación de riesgo*. La otra, por la *situación de éxito* que ha sido capaz de alcanzar el individuo. En cuanto a las *situaciones de riesgo*, y aunque aquí no vamos a desarrollar en profundidad el tema de la definición de este tipo de situaciones, sí hay que tener presente que pueden configurarse a través de la confluencia de un conjunto de factores —los *factores de riesgo*— que pueden proceder tanto del propio individuo como su entorno y de las relaciones recíprocas entre ambos. En cuanto a la *situación de éxito*, normalmente se utiliza el criterio de un resultado que implique un nivel de logro, por ejemplo, el hecho de tener un empleo y sentirse satisfecho de su vida (Spekman, Herman y Vogel, 1993)⁶, tener un buen rendimiento escolar o un buen progreso en la escuela (en el caso de los niños y adolescentes) (Spekman, Herman y Vogel, 1993)⁷ e incluso no ser distintos de la norma en dimensiones de escuela/trabajo, relaciones con la pareja, relaciones con los hijos, relaciones con el grupo de iguales y en cuanto a las autovaloraciones de sí mismos (Werner, 1993)⁸.

Parece claro que la noción de *resilience* emerge al enfocar los aspectos positivos de la vida del niño más que los aspectos negativos como ocurre cuando se trabaja desde la perspectiva de los factores de riesgo.

El estudio de este fenómeno conduce a la búsqueda de mecanismos o factores que han contribuido a que determinados individuos en situación de riesgo por razón de sus condiciones familiares y sociales adversas hayan podido hacerles frente y llegar a

4 Citado por Casas (1994). Previsió i qualitat de vida dels infants. *Revista d'informació i investigació socials*, 0, p. 43.

5 Citado por Miller, M. (1995). Op. cit. p. 3.

6 Citado por Miller, M. (1995). Op. cit. p. 4.

7 Citado por Miller, M. (1995). Op. cit. p. 4.

8 Citado por Miller, M. (1995). Op. cit. p. 4.

alcanzar un cierto éxito en las distintas esferas de su vida. Estos factores se denominan *factores protectores*. Para Manciaux (1994) los factores protectores son los que ayudarían a los individuos a enfrentarse y superar las adversidades. Clark (1995:4) define los factores protectores como «salvaguardas individuales o ambientales que potencian las habilidades de los adolescentes para resistir los acontecimientos estresantes de la vida y, a la vez, les ayuda a adaptarse a la situación y a desarrollar competencias para enfrentarse a estas situaciones».

De hecho, los factores protectores serían opuestos a los factores de riesgo. Así, mientras que los factores de riesgo contribuyen a aumentar la probabilidad de que un determinado problema se desarrolle⁹ los factores protectores tienden a disminuir dicha probabilidad. Sin embargo, la diferencia entre uno y otro tipo de factores radica en que los factores protectores operan solamente cuando el riesgo está presente, es decir, los factores protectores siempre deben vincularse a la presencia de riesgos. Refiriéndonos al riesgo de fracaso escolar¹⁰ Pianta (1990) define los factores protectores como aquellos que distinguen individuos de alto riesgo que obtienen unos buenos resultados escolares de aquellos que, tal como se había predicho, presentan problemas. Del mismo modo que los factores de riesgo, pueden proceder del propio individuo, de las relaciones con otras personas y del entorno social.

Desde esta perspectiva, *prevenir* supone realizar intervenciones cuya consecuencia sea la eliminación de factores de riesgo o —más probablemente— la disminución de sus efectos, y la potenciación de los factores protectores. Por lo tanto, prevenir significa actuar sobre factores que sean modificables. Prevenir el fracaso escolar implica actuar sobre factores de riesgo y factores protectores que sean modificables. Si nos situamos en la escuela, la intervención deberá intentar incidir en factores modificables a través de la intervención educativa que se lleva a cabo en ella. Si nos situamos en el marco comunitario, se deberá incidir en factores modificables mediante la intervención educativa y social desarrolladas en este marco. Es evidente la importancia de la coordinación entre ambos para una mejor prevención del fracaso escolar.

En EE.UU., con una tradición de dos décadas de investigación sobre este tema, se habla de factores protectores como variables que conciernen a las *características y*

9 En general se considera que los *factores de riesgo* poseen una asociación causal con un determinado problema. Si nos referimos al fracaso escolar, los factores de riesgo serán factores personales, familiares, sociales o escolares que tienen algún tipo de asociación causal con el desarrollo del fracaso escolar, es decir, contribuyen a aumentar la probabilidad de que éste se desarrolle. No obstante, no tienen que se necesariamente causas directas del problema. Un factor de riesgo aumenta la probabilidad de que se desarrolle el problema pero NO conduce a él irremisiblemente.

10 Tampoco puedo aquí extenderme en la explicación del significado de riesgo. Sin embargo cabe señalar que en todos estos temas hay una premisa previa que es *la definición del problema en relación al cual se habla de riesgo*. Si el problema es el fracaso en la escuela podremos encontrar factores de riesgo y factores protectores del desarrollo de dicho problema. Si el problema es el consumo de drogas en la adolescencia quizá hallemos algunos factores comunes pero también es posible que otros factores tomen relevancia o que aparezcan factores de riesgo o protectores distintos. Uno de los principales problemas que encontramos en la literatura en cuanto al uso del término riesgo es que con frecuencia no se manifiesta explícitamente cuál es el problema en cuestión. Así, se habla de factores de riesgo pero en ocasiones nos encontramos con la indefinición de la problemática, con lo que el término llega a ser ambiguo y equívoco.

a las habilidades que poseen los jóvenes «resilientes», así como a las *variables ambientales que influyen positivamente en éstas*. Parece que en estos casos el niño tiene un conjunto o un patrón de factores protectores en lugar de tener solamente uno o dos (Guetzloe, 1994).

3. REVISIÓN DE INVESTIGACIONES DIRIGIDAS A IDENTIFICAR FACTORES PROTECTORES

Winfield (1996) distingue entre dos conceptos de *resilience*, con repercusiones en el campo de la investigación. El primero supone entender este fenómeno como un *atributo de las personas*. Es un concepto estático. Ser una persona «resiliente» implica tener una serie de características. El segundo implica entender este fenómeno como *mecanismos protectores* que modifican las respuestas individuales en situaciones de riesgo. Es una visión dinámica. A nuestro entender ambas posiciones no son excluyentes.

Las investigaciones que se han llevado a cabo tienden, en general, a buscar *características de los individuos*. En este sentido, éstas considerarían la *resilience* como atributo. Sin embargo, la mayoría, además de buscar estos rasgos del individuo, han tratado de hallar *características del entorno social y familiar* que pueden haber ejercido la función de factores protectores en personas en situación de riesgo. Hay que suponer que el interés por identificar estos rasgos se debe al intento de encontrar aspectos sobre los que incidir y en los que basar el diseño de intervenciones orientadas a promover en los individuos la capacidad para resistir ante circunstancias adversas.

En las tablas 1 y 2 figuran los factores considerados por varios autores como factores protectores. Todos se basan en investigaciones sobre individuos «resilientes». Hemos clasificado dichos factores en dos grupos. En primer lugar, los que se refieren a *características personales* de los individuos, entre los que hemos incluido los aspectos actitudinales, los aspectos cognitivos e intelectuales, los rasgos afectivos y de personalidad y las habilidades sociales y de interacción social (Tabla 1). En segundo lugar, los que se refieren a *características del entorno familiar y social*, que hemos podido agrupar en cuatro bloques: las características familiares, tanto de tipo estructural como de tipo dinámico, las características de la comunidad y del entorno social, las características escolares, y un último grupo que, según señalan varios autores, está relacionado con el hecho de haber podido establecer relaciones sociales significativas con personas concretas a lo largo de la vida (Tabla 2).

Sintetizando las aportaciones de los estudios, cabe indicar que entre los *aspectos actitudinales* se encuentra el tener una autoestima positiva, valorarse a sí mismos, tener una visión positiva de sus posibilidades y de su futuro, poseer un elevado grado de autonomía, un sentido de control de su vida, de sus éxitos y de sus fracasos y considerar la escuela como una experiencia valiosa. Entre los *aspectos cognitivos* se puede citar el tener habilidades para la resolución de problemas y para tomar sus decisiones. Entre los *aspectos afectivos* destacan el caracterizarse por una cierta estabilidad emocional y tener unas características temperamentales que producen respuestas sociales positivas por parte de los demás. Estas respuestas positivas también vienen facilitadas

Tabla 1
CARACTERÍSTICAS PERSONALES VINCULADAS AL FENÓMENO
DE LA «RESILIENCIA»

AUTORES	Aspectos actitudinales	Aspectos cognitivos e intelectuales	Rasgos afectivos y de personalidad	Habilidades sociales y tipo de interacciones establecidas
Garmezy (1983, citado por Winfield, 1996)	<ul style="list-style-type: none"> • Autoestima positiva • Sentido de poder personal • <i>Locus of control</i> interno 	<ul style="list-style-type: none"> • Inteligencia • Habilidades críticas de solución de problemas 	<ul style="list-style-type: none"> • Sentido del humor • Bajo nivel de destructividad y agresividad • Estabilidad emocional 	<ul style="list-style-type: none"> • Poseer habilidades sociales • Alto grado de sensibilidad social • Empatía • Alto nivel de cooperación y participación • Interacciones positivas con el grupo de iguales y con los adultos
Connell, J.P.; Spencer, M.B.; Aber, J.L. (1994)	<ul style="list-style-type: none"> • Sentimiento de autovaloración • Sentido de control sobre sus éxitos y sus fracasos en la escuela 		<ul style="list-style-type: none"> • Seguridad emocional 	
Miller, M. (1995) (Trabaja con jóvenes con dificultades de aprendizaje)	<ul style="list-style-type: none"> • Autodeterminación • Reconocimiento de la existencia de la dificultad de aprendizaje 	<ul style="list-style-type: none"> • Tener algunas áreas académicas especialmente fuertes 		
Werner y Smith (1989) citado por Burger (1994)	<ul style="list-style-type: none"> • Autoestima • Eficiencia personal, capaces de pedir ayuda cuando la necesitan • Considerar la escuela como una experiencia valiosa 	<ul style="list-style-type: none"> • Concentración y capacidad de atención • Habilidades para resolver problemas • Habilidades para tomar decisiones 	<ul style="list-style-type: none"> • Características temperamentales que potencian respuestas sociales positivas por parte de los padres, compañeros y profesores 	<ul style="list-style-type: none"> • Gusto por compartir las cosas y por realizar actividades en grupo
Berliner, B.; Benard, B. (1995)	<ul style="list-style-type: none"> • Autonomía: habilidad para actuar independientemente y ejercer control sobre el propio entorno • Tener sentido de la propia identidad • Sentido de futuro: habilidad para imaginarse un futuro propio, ser optimista y aspirar a un buen rendimiento personal y educativo 	<ul style="list-style-type: none"> • Habilidad para tomar decisiones de forma creativa, crítica y reflexionada • Habilidad para buscar ayuda de los demás, reconocer vías alternativas para solucionar problemas y resolver conflictos 	<ul style="list-style-type: none"> • Mantener el sentido del humor 	<ul style="list-style-type: none"> • Competencia social o habilidad para establecer y sostener relaciones positivas • Empatía • Relaciones que producen confianza y respeto
Clark, P. (1995)	<ul style="list-style-type: none"> • Autoestima y autoeficacia • Responsabilidad personal • Alto nivel de optimismo • Percepción de que las experiencias son constructivas • Visión positiva del futuro personal • Habilidades para enfrentarse a las situaciones • Aspectos actitudinales relacionados con la escuela: motivación, aspiraciones positivas, buenas ejecuciones escolares, implicado en actividades extracurriculares 	<ul style="list-style-type: none"> • Habilidades intelectuales • Habilidades de solución de problemas • Habilidades de toma de decisiones 		<ul style="list-style-type: none"> • Habilidades sociales e interpersonales • Otorgar valor al hecho de ayudar a la gente • Habilidad para ganar la atención positiva de los demás • Tener cuidado con los sentimientos de los demás • Habilidades para hacer amigos • Influencias positivas del grupo de iguales
Organización Mundial de la Salud, citado por Manciaux (1994)	<ul style="list-style-type: none"> • Autoestima • Autonomía • Desarrollo de una imagen positiva de sí mismo/a • Sentido de eficacia y recursos personales 			

Tabla 2
CARACTERÍSTICAS FAMILIARES Y SOCIALES VINCULADAS AL FENÓMENO DE LA «RESILIENCE»

AUTORES	Características familiares	Tener personas significativas a lo largo de su vida	Características de la comunidad, entorno social	Características escolares
Garnezy (1983, citado por Winfield, 1996)	<ul style="list-style-type: none"> • Los padres se preocupan por la educación de sus hijos y participan de esta educación • Los padres dirigen las tareas de sus hijos diariamente • Los padres son conscientes de los intereses y de los objetivos de los niños y niñas 	<ul style="list-style-type: none"> • Los jóvenes tienen, como mínimo, un adulto significativo en sus vidas 		
Connell, J.P.; Spencer, M.B.; Aber, J.L. (1994)	<ul style="list-style-type: none"> • Apoyo de las familias 			
Miller, M. (1995) (Trabaja con jóvenes con dificultades de aprendizaje)		<ul style="list-style-type: none"> • Haber tenido un/a profesor/a que les ofreció apoyo • Tener un amigo especial 		
Wyman et al, 1991, citado por Miller, (1995)	<ul style="list-style-type: none"> • Disponibilidad de apoyo significativo por parte de la familia extensa y otros adultos 			
Werner y Smith (1989), citado por Burger (1994)	<ul style="list-style-type: none"> • Familias de menos de 4 hijos • Ser hijos mayores que comparten responsabilidades familiares • Normas familiares claras 	<ul style="list-style-type: none"> • Red de adultos de los que como mínimo uno representaba para el/la joven un modelo a seguir • Tener un amigo íntimo 	<ul style="list-style-type: none"> • Haber tenido «segundas oportunidades en la sociedad (tanto en la escuela, el trabajo, etc.) que capacitan a los jóvenes en alto riesgo para adquirir competencia y confianza. 	
Berliner, B.; Benard, B. (1995)	<ul style="list-style-type: none"> • Expectativas positivas que se comunican a los hijos 		<ul style="list-style-type: none"> • Oportunidades de participar en actividades a que son consideradas de valor por parte de la familia la escuela y la comunidad en general 	
Clark, P. (1995)	<ul style="list-style-type: none"> • Apoyo familiar • Disciplina de los padres • Supervisión de los padres • Comunicación con los padres 	<ul style="list-style-type: none"> • Tener un amigo próximo • Relaciones estrechas con una persona de la familia como mínimo 	<ul style="list-style-type: none"> • Pertener a una comunidad que le ofrece apoyo • Disponer de otros recursos adultos y comunicación con ellos • Estar implicado en organizaciones comunitarias 	<ul style="list-style-type: none"> • Clima positivo en la escuela
Organización Mundial de la Salud, citado por Manciaux (1994)	<ul style="list-style-type: none"> • Un entorno familiar (y/o substitutivo) que pueda responder a las necesidades de desarrollo del niño • Situación económica favorable 		<ul style="list-style-type: none"> • Buena integración en la comunidad 	<ul style="list-style-type: none"> • Buena integración en la escuela

por el hecho de que estas personas *poseen habilidades para la relación interpersonal* y empatía de modo que son capaces de establecer relaciones positivas tanto con el grupo de iguales como con los adultos.

En cuanto a las *características familiares*, éstas se pueden resumir en el hecho de tener un entorno familiar que les apoya en el sentido de tener unas actitudes y unas expectativas positivas de los padres hacia los hijos, así como una dirección y supervisión por parte de los padres. En cuanto al *entorno social* más amplio destaca el hecho de haber tenido a lo largo de su infancia y adolescencia adultos significativos (distintos de los padres) y algún amigo o amiga especial. También parece que el apoyo que puede proporcionar la comunidad a través de sus recursos puede ser un factor que contribuya a la «resiliencia» de las personas.

Entre los estudios revisados que han realizado sus trabajos con personas «resilientes» sólo se destacan dos aspectos relacionados con el entorno escolar: el clima positivo y la buena integración del niño o la niña en este medio.

La investigación que presentamos a continuación se sitúa en la línea de los estudios sobre personas, adolescentes en este caso, que han logrado un cierto nivel de éxito escolar a pesar de las condiciones sociales adversas. Nos interesaba sobre todo intentar identificar factores protectores que pudieran ser aspectos modificables mediante la intervención educativa que se lleva a cabo en la escuela. Nos pareció que podía ser interesante esta perspectiva por cuanto ofrecía elementos útiles para articular intervenciones educativas con carácter preventivo.

4. UN ESTUDIO DE CASOS PARA LA IDENTIFICACIÓN DE FACTORES PROTECTORES DEL RIESGO DE FRACASO ESCOLAR

4.1. Objetivo

El objetivo de la investigación que se presenta era identificar algunas variables que habían podido ejercer la función de factores protectores del fracaso escolar. Entendiendo, como hemos dicho antes, que un factor protector es una variable que contribuye a disminuir el riesgo de fracaso escolar en individuos en situación de riesgo, era necesario que nuestra investigación se centrara en el estudio de chicos y chicas «resilientes» en el sentido de haber conseguido un cierto nivel de logro escolar a pesar de haber vivido circunstancias adversas y desfavorables al éxito en la escuela. Esto permitiría conocer qué aspectos habrían ayudado a compensar el riesgo.

Encontrar factores protectores nos permitiría hallar algunas directrices para intervenir educativamente en la escuela. Estas líneas serían útiles no solamente para los niños y niñas en situación de riesgo sino para todos los escolares en general.

4.2. La selección de los casos

Basándonos en la definición del fenómeno de la «resilience» que hemos visto anteriormente, en nuestra investigación un individuo «resiliente» tenía que ser entendido como aquél que había alcanzado un cierto nivel de éxito en la escuela a pesar de hallar-

se en situación de riesgo. Por lo tanto, debíamos conseguir identificar algunos niños y niñas que se encontraran en situación de riesgo de fracaso escolar y, sin embargo, hubieran obtenido unos resultados escolares satisfactorios.

Así, la situación de riesgo se definió como el hecho de vivir separados de sus familias, en comunidades infantiles dependientes de la Dirección General de Atención a la Infancia de la Generalitat de Catalunya¹¹. Aunque no consideramos relevante para nuestra investigación conocer qué hecho o hechos habían conducido a la separación de estos jóvenes de sus familias en un momento dado de su vida, el hecho de residir en centros residenciales era un indicador claro de la pertenencia de estos chicos y chicas a un grupo de alto riesgo de fracaso escolar, con problemas familiares suficientemente graves para motivar la separación del hijo o la hija de su familia. Además, las investigaciones sobre el rendimiento escolar indicaban claramente que el hecho de tener un origen social desfavorecido y de haber vivido situaciones de desestructuración familiar podía constituir un factor de riesgo importante asociado al fracaso escolar. Aunque somos conscientes de que el fenómeno de la «resiliencia» no se reduce al hecho de tener éxito en la escuela y obtener unas buenas calificaciones —quizá ni siquiera sea el hecho más importante si tomamos en consideración la persona en toda su complejidad— en términos de nuestra investigación constituía un indicador que podíamos utilizar como criterio de selección. Además, teniendo presente que nuestro objetivo era identificar variables relevantes para la prevención del fracaso escolar, era lógico que a pesar de las limitaciones utilizáramos como signo de «*resilience*» algún indicador relacionado con el éxito en la escuela secundaria. De algún modo estos jóvenes habían logrado hasta el momento hacer frente a las condiciones desfavorables que los hacían vulnerables al fracaso escolar, siendo capaces de hallar sus propios recursos para situarse ante la realidad escolar y ante el aprendizaje.

Centramos el estudio en la provincia de Girona. De los chicos y chicas que el año 1993 se encontraban bajo la tutela de la Dirección General de Atención a la Infancia, seleccionamos aquéllos que presentaban un cierto nivel de éxito escolar, determinado mediante diferentes indicadores:

1. Haber superado la EGB y estar realizando Formación Profesional, ya que ninguno de ellos estaba estudiando BUP en esos momentos.
2. Estar matriculado y asistir a un centro de educación secundaria en el momento de realizar esta investigación.
3. Haber obtenido una calificaciones escolares, especialmente durante el curso 1992-93, que supusieran haber superado, en términos de calificaciones, la mayoría de las asignaturas.

11 En Cataluña reciben el nombre de Centros Residenciales de Acción Educativa según la Orden del 15 de julio de 1987 (DOGC. núm. 878, de 17.8.87) que establece la tipología de entidades, servicios y establecimientos sociales.

Partiendo de todos estos indicadores obtuvimos un primer listado de 9 chicas y 6 chicos. Una vez realizadas unas primeras entrevistas a los chicos y a sus educadores en el centro residencial pudimos encontrar cuáles eran los casos que más se ajustaban a los criterios de joven «resiliente». Así, el estudio definitivo se llevó a cabo con 5 casos, tres chicas y dos chicos.

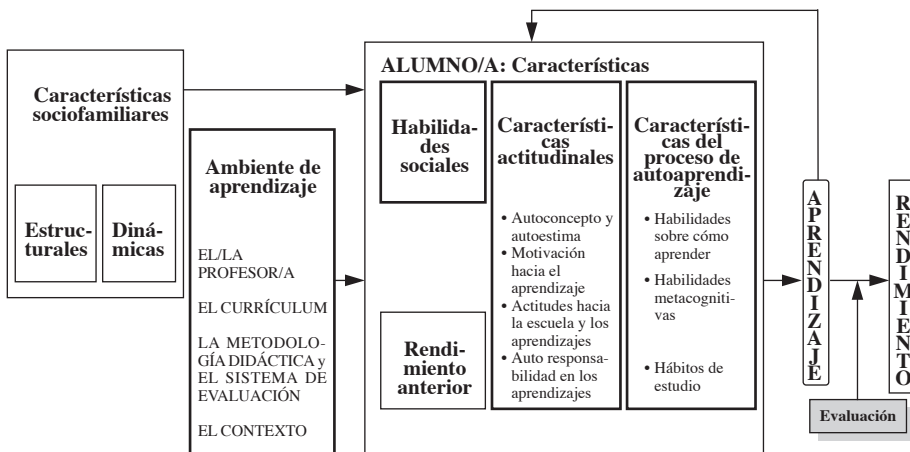
4.3. Los aspectos estudiados

Para seleccionar los aspectos a estudiar partimos de un modelo que denominamos «Modelo de Riesgo». Este modelo se elaboró a partir del análisis de investigaciones sobre rendimiento escolar, y a partir del estudio del concepto de riesgo y de las investigaciones sobre riesgo¹². El modelo (Figura 1) describe las relaciones entre variables que son modificables a través de la intervención educativa escolar y que pueden estar relacionadas con el aumento o la disminución de la probabilidad del riesgo de fracaso escolar. Su configuración se basa en la consideración de que los estudiantes tienen un papel activo en su aprendizaje.

Figura 1

MODELO DE ANÁLISIS DE LAS VARIABLES QUE INFLUYEN EN EL RIESGO DE FRACASO ESCOLAR. VARIABLES Y RELACIONES

(Los recuadros con línea gruesa indican los conjuntos de variables modificables mediante la intervención educativa escolar. Las flechas indican la relación entre las distintas variables y los aprendizajes)



12 Para más detalles sobre la elaboración de este modelo ver Fullana Noell, J. (1996). La prevención del fracaso escolar: Un modelo para analizar las variables que influyen en el riesgo de fracaso escolar. *Bordón*, 48(2). 151-165.

No todas las variables de modelo fueron estudiadas en esta investigación. Optamos por seleccionar aquéllas relacionadas con las maneras con que cada sujeto hace frente a los aprendizajes y al contexto donde éstos se desarrollan, es decir, orientamos nuestra investigación sobre el estudio de las características actitudinales de los estudiantes más vinculadas a los aprendizajes (Tabla 3), las características de su proceso de aprendizaje (Tabla 4) y algunos rasgos de su entorno de aprendizaje (Tabla 5). En las tablas 3, 4 y 5 se presenta una breve definición de los aspectos que nos interesaba conocer sobre cada variable.

Además de los aspectos mencionados en las tablas, se obtuvo información sobre *la valoración que los propios sujetos hacían de los resultados académicos* del curso que estaban estudiando y *la valoración sobre los aprendizajes adquiridos* en función de su utilidad futura y en función de la importancia percibida para el crecimiento personal.

En cuanto al entorno familiar, únicamente se tuvieron en cuenta aquellos aspectos que eran percibidos por el propio sujeto como variables relevantes para explicar su éxito escolar.

Tabla 3

VARIABLES RELACIONADAS CON LAS CARACTERÍSTICAS ACTITUDINALES DE LOS/LAS ESTUDIANTES

VARIABLES	DEFINICIÓN
Autoconcepto-autoestima (Frey, 1989; González y Tourón, 1992)	Nos interesa saber como el/la joven se concibe a sí mismo como estudiante: qué pensamientos, qué sentimientos y qué actitudes cree que son representativas de sí mismo como estudiante. La valoración que la persona hace de sí misma, de sus propias características como estudiante forma parte de la autoestima del sujeto.
Motivación hacia la escuela y el trabajo escolar (Ball, 1988)	Orientación y mantenimiento de las conductas hacia la realización de tareas relacionadas con la escuela y el trabajo escolar.
Actitudes hacia la escuela y el trabajo escolar (Ball, 1988)	Aquellos sentimientos o tendencias favorables o contrarias a la escuela en general y al trabajo escolar que en ella se realiza.
Autoresponsabilidad en el aprendizaje (locus of control, motivación del logro y atribución de la causalidad) (Wang y Peeverly, 1987)	En nuestro estudio consideraremos que el individuo tiende a un locus de control interno: <ul style="list-style-type: none"> • Si se manifiesta en el sentido que él cree que es determinante de los acontecimientos de su entorno, concretamente si considera que es él o ella misma la determinante de sus resultados académicos, y no los atribuye a la suerte o a las circunstancias. • Si se considera responsable de sus conductas relacionadas con la adquisición de aprendizajes académicos y de sus propios resultados. • Si atribuye sus éxitos o fracasos a variables que él o ella pueden controlar, como el esfuerzo, y no los atribuye a la falta de habilidades o capacidades. La motivación del logro es, según Ball, 1988: <ul style="list-style-type: none"> • La predisposición que tiene el chico o la chica ante el éxito en la escuela. • Su disposición ante la posibilidad de conseguir unos mejores resultados escolares, la orientación y empeño del sujeto hacia la realización correcta de las tareas escolares. • Su orientación hacia el logro lo que implica un cierto grado de perseverancia en la realización de una tarea.
Rendimiento anterior	Resultados del alumno o de la alumna en cursos anteriores, de una manera general, por ejemplo, si obtuvo el graduado escolar y con qué calificación, las asignaturas aprobadas y suspendidas, o bien las calificaciones de las distintas asignaturas que figuran en las actas académicas.

Tabla 4

VARIABLES RELACIONADAS CON LAS CARACTERÍSTICAS DEL PROCESO DE APRENDIZAJE DE LOS/LAS ESTUDIANTES

VARIABLES	DEFINICIÓN
Hábitos de estudio (Rodríguez Espinar, Álvarez et al, 1993)	<p>Son habilidades conductuales que incluyen las habilidades para la planificación del estudio y las habilidades de autocontrol como la concentración y reducción de la ansiedad, las habilidades de control del ambiente y las habilidades de control del progreso.</p> <p>Dentro de la categoría de hábitos de estudio hemos tomado en consideración para nuestro estudio:</p> <ul style="list-style-type: none"> • La constancia en el trabajo. • el tiempo de dedicación al trabajo relacionado con las tareas escolares, • la planificación en conjunto del trabajo a realizar, • la tendencia del alumno a buscar explicaciones del profesor sobre los resultados de la tareas con el fin de realizar una autoevaluación.
Habilidades de autoaprendizaje (Wang y Pevery, 1987)	<p><i>Habilidades sobre como aprender</i></p> <p>El conocimiento que los estudiantes tienen sobre estrategias que faciliten la codificación, el recuerdo, la comprensión y la asimilación de nuevos conocimientos. Entre estas habilidades: hacer guiones, resúmenes, esquemas, saber reconocer cuando se necesita ayuda y saber como buscarla y donde hallarla.</p>
	<p><i>Habilidades metacognitivas</i></p> <p>Habilidades de identificación de las demandas, de planificación de una tarea, entendida como el proceso por el cual un individuo estructura la tarea en función de los objetivos; y la autosupervisión de la tarea, referida a la evaluación continuada de la tarea lo cual supone adecuar el proceso de elaboración a los objetivos establecidos, modificando de forma flexible las planificaciones previas y las estrategias utilizadas.</p>
Reflexividad/impulsividad (Forns, 1989)	<p>Entenderemos que un/a estudiante tenderá a ser impulsivo/a si ante una tarea tiene tendencia a pasar a la acción sin pararse a pensar, mientras que será de tendencia reflexiva si ante una tarea pasa a la acción una vez transcurrido un tiempo más o menos prolongado de reflexión.</p>

Tabla 5

VARIABLES RELACIONADAS CON LAS CARACTERÍSTICAS DEL AMBIENTE DE APRENDIZAJE

VARIABLES	DEFINICIÓN
Los profesores/as	<ul style="list-style-type: none"> • Concepto que los profesores tienen del alumno/a alumna como estudiante • Expectativas de los profesores en cuanto a los estudios del alumno o de la alumna • Características de la relación del alumno con los profesores • Estilo docente, entendido como las conductas referidas a la enseñanza
El currículum	<ul style="list-style-type: none"> • Asignaturas que el alumno prefiere • Organización del trabajo en el aula: tipo de trabajo que prefiere el estudiante, trabajo en equipo o trabajo individual • Tareas y actividades que el/la estudiante prefiere • Tipo de tareas que el /la estudiante considera que le son de mayor utilidad para aprender
El contexto de aprendizaje	<ul style="list-style-type: none"> • Características de los estudiantes que forman el grupo clase según la percepción del sujeto • La relación del chico o la chica con sus compañeros de clase

4.4. Método

El diseño utilizado para llevar a cabo este estudio está en consonancia con los denominados *estudios de casos*. En coherencia con nuestro punto de partida, es decir, con el hecho de entender que los estudiantes tienen un papel activo en su aprendizaje, en considerarlos mediadores entre su aprendizaje y su entorno (Wang y Peverly, 1987), optamos por hacer un estudio de casos de tipo *cualitativo* que nos permitiera captar las percepciones que los estudiantes tenían de la escuela, del ambiente de aprendizaje, del propio proceso de aprendizaje, de sí mismos como estudiantes. Además, el que se tratara de estudiar casos atípicos hizo que consideráramos más adecuado realizar un estudio de casos de corte cualitativo. Optar por un estudio de casos implica, según Yin (1990), investigar un fenómeno contemporáneo dentro de su contexto de la vida real y utilizando múltiples fuentes de evidencia. En nuestra investigación, más que estudiar al individuo dentro de su contexto estudiamos cómo percibía dicho contexto, cómo se veía a sí mismo dentro de este contexto, qué aspectos de su vida cotidiana —relacionados con la familia, el centro residencia y el centro docente— consideraba más relevantes para explicar su propio desarrollo escolar. Nos interesaba conocer las vivencias de los sujetos, aprender el significado que para ellos tenía su experiencia escolar y qué aspectos eran valorados por ellos mismos como significativos para ayudarlos a enfrentarse a su situación y a alcanzar unos logros escolares satisfactorios.

Para conseguir nuestro objetivo decidimos que el instrumento de recogida de datos que más se ajustaba a nuestro propósito era la *entrevista*. El hecho de utilizar entrevistas derivaba directamente de nuestro objetivo de recoger información a partir de las percepciones de los chicos y chicas sobre como se sitúan ante sus aprendizajes, complementando esta visión con la de sus educadores y educadoras con la función de tutores de los chicos, y de sus profesores y profesoras tutores en el instituto. Como hemos dicho, queríamos conocer cómo percibían los sujetos los diferentes temas más que obtener una medida objetiva de las diferentes variables.

Por ello diseñamos entrevistas semiestructuradas, no estandarizadas (Goetz y LeCompte, 1988)¹³, que nos permitieron obtener datos sobre las propias y vivencias. La entrevista cualitativa resultaba adecuada ya que, como señalan Schatzman y Strauss (1973)¹⁴ su objetivo fundamental es obtener datos reveladores de los significados de los participantes, mostrar como éstos conciben sus mundos y como explican estas concepciones. Las entrevistas a los jóvenes, complementadas con las entrevistas a los profesores y educadores nos permitió triangular la información y componernos una imagen comprensiva y global de cada uno de los casos en relación a los aspectos objeto de nuestro interés.

13 La entrevista no estandarizada es para Goetz y LeCompte (1988) una guía en la cual se anticipan las cuestiones generales y la información específica que el investigador quiere obtener. Su enfoque es informal y ni el orden de las preguntas ni su contexto están prefijados.

14 Citado por Goetz, J.P.; LeCompte, M.C. (1988). *Etnografía y diseño cualitativo en investigación educativa*, Madrid: Morata.

Para cada uno de los cinco casos obtuvimos información del propio sujeto, de su profesor tutor de curso y de su educador tutor en el centro residencial. Las entrevistas fueron complementadas con unas escalas adaptadas con el objetivo de obtener información complementaria en torno a habilidades metacognitivas y la autoresponsabilidad en el aprendizaje (Zimmerman y Martínez-Pons, 1986, 1988). Una de estas escalas iba dirigida a los estudiantes y la segunda a los profesores y a las profesoras.

El *análisis de los datos* se llevó a cabo mediante técnicas de *análisis de contenido* (Bardin, 1986). En primer lugar se elaboró un informe sobre cada uno de los casos y finalmente se procedió a su comparación para cada una de las dimensiones estudiadas: las características actitudinales de los estudiantes, las características de su proceso de aprendizaje, los rasgos del ambiente de aprendizaje, los resultados y las características del entorno vinculadas con la motivación para los estudios. A partir de esta comparación obtuvimos algunos indicios sobre variables que podían haber resultado relevantes como factores protectores en estos casos de alto riesgo.

4.5. Los resultados

El estudio de cada uno de los casos y su comparación aporta elementos que nos ayudan a comprender por qué estos estudiantes, que han vivido en condiciones familiares y sociales desfavorables, han logrado un nivel considerable de éxito escolar. La comparación de los casos nos permitió observar algunos aspectos que podían haber actuado como factores protectores del riesgo de fracaso escolar, es decir, como factores que habrían contribuido a hacer a estos jóvenes menos vulnerables a la influencia de los factores de riesgo derivados principalmente de las condiciones socio-familiares.

En cuanto a las *características actitudinales de estos jóvenes más vinculadas a los aprendizajes* observamos lo siguiente:

1. Todos los chicos y chicas tenían un *concepto positivo de sí mismos*. No todos se consideraban buenos estudiantes aunque sí eran capaces de hallar aspectos positivos de sí mismos como la imaginación, la memoria, la capacidad para estudiar. Incluso alguno de ellos tenía un concepto excesivamente elevado de sí mismo según apreciaban sus educadores y profesores. Éstos valoraban positivamente las capacidades y aptitudes de los jóvenes para sacar adelante sus estudios.
2. Manifestaban una *actitud positiva hacia el instituto y, especialmente, hacia el aprendizaje*. Aunque esta actitud variaba en intensidad, en ninguno de los casos se evidenció ningún tipo de rechazo hacia todo lo que conllevan los estudios. Al contrario, tendían a ver en sus estudios una vía que les podía facilitar el acceso a un puesto de trabajo y, como consecuencia, y desde su punto de vista, el acceso a una situación social estable y normalizada.
3. Se mostraban *motivados hacia el estudio* aunque no lo estaban para realizar todos los tipos de tareas que se les exigían. A pesar de que el nivel de motivación no era el mismo en todos ellos, cada uno se había propuesto un objetivo a medio o largo plazo que les motivaba y les predisponía a aceptar y aprovechar lo que el sistema educativo formal les estaba ofreciendo. Esta meta era diferente en cada

joven. Los objetivos variaban desde el más pragmático como conseguir un título con el que acceder a un puesto de trabajo, hasta el de considerar que los estudios tienen sentido en sí mismos, como medio para el crecimiento intelectual y cultural, y para obtener los conocimientos y habilidades básicas que les habrán de permitir acceder —si las condiciones lo favorecían— a otros niveles superiores del sistema educativo. Por lo tanto, la motivación de estas tres chicas y dos chicos no se centraba en la realización de las tareas concretas sino en la meta que, con más o menos claridad, más explícitamente o menos, se habían marcado a medio o a largo plazo.

4. En lo referente a la *autoresponsabilidad en el aprendizaje* hemos constatado una mayor variación entre los casos. Dos de las chicas de entre los cinco casos estudiados aparecieron como las más autoresponsables en su aprendizaje. A la vez, es en ellas en las que se vio con más claridad que habían logrado un nivel de éxito escolar importante ya que estaban cursando el segundo ciclo de Formación Profesional. Ambas parecían tener una elevada *motivación del logro*, es decir, se mostraban predispuestas al éxito, se esforzaban en realizar las tareas lo mejor posible, en conseguir unos resultados escolares mejores, a pesar de que en ocasiones no lo conseguían. A las dos les agradaba saber qué errores habían cometido con el fin de hacer sus tareas y ejercicios mejor en una próxima ocasión. Manifestaban con más claridad que los demás *un locus de control interno*. Se consideraban a sí mismas responsables de sus resultados académicos, de que era su esfuerzo y no la suerte o las circunstancias, el que estaba directamente relacionado con sus resultados académicos y su tendencia era *atribuir sus éxitos y sus fracasos a variables que podían controlar*. Los tres chicos restantes mostraban también motivación para superar el curso pero la motivación del logro no resultaba tan evidente y pareció que se movían más para cumplir el mínimo que les permitiese ir superando las distintas tareas.

En cuanto a las *características del proceso de autoaprendizaje* hay que señalar que éstas no destacaron de forma especial como posibles factores protectores:

1. En relación a las *habilidades de autoaprendizaje* las únicas estrategias que utilizaban fueron las de hacer resúmenes y tomar apuntes, para ayudarse a sí mismos a aprender los contenidos de las asignaturas. No obstante, estas estrategias formaban parte del tipo de tareas que se les exigía desde el instituto, por lo que hubiera resultado arriesgado admitir que se utilizaban espontáneamente como medios para mejorar el aprendizaje.
2. En lo referente a las *habilidades metacognitivas*, sólo algunas habilidades de planificación de las tareas y de autosupervisión pareció que se activaban cuando la tarea resultaba interesante y muy motivadora para los chicos y las chicas.

Aunque estos dos puntos destacaron como características del proceso de aprendizaje, debemos reconocer que las entrevistas y las escalas no resultaron suficientes para poder establecer con precisión qué estrategias utilizaban y qué habilidades metacog-

nitivas poseían y utilizaban de forma espontánea. Lo que quedó bastante claro fue que los chicos y las chicas tenían dificultades para expresar qué estrategias ponían en funcionamiento para llevar a cabo los diferentes tipos de tareas académicas y para aprender.

Tampoco el tema de los *hábitos de estudio* destacó especialmente como posible factor compensador del riesgo, a pesar de que la información obtenida apuntaba hacia una mayor relevancia de estos hábitos que de las habilidades de aprendizaje. Dos de las chicas fueron las que habían adquirido unos ciertos hábitos como la *constancia en el trabajo*, de *saber solicitar ayuda* del profesor cuando lo necesitaban, así como de *organización del estudio*, para poder compaginar estudio y empleo. Los otros tres casos se orientaban en función de las demandas, es decir, realizaban las tareas según las iban exigiendo los profesores y profesoras como parte del trabajo académico.

En cuanto al *ambiente de aprendizaje* debemos señalar que:

1. Tanto los *profesores como los educadores tenían un buen concepto de estos jóvenes como estudiantes*. Los consideraban personas con capacidades y aptitudes para el estudio. También sus *expectativas hacia ellos eran positivas* aunque todos tenían presente los condicionantes que los chicos y las chicas no podían controlar como es el hecho de tener que independizarse en el momento de cumplir los dieciocho años¹⁵. Además, los jóvenes conocían cual era el concepto que de ellos tenían los adultos. Sabían qué pensaban sus profesores/as y sus educadores/as y sabían que la imagen que los adultos se habían forjado de ellos era bastante positiva.
2. No tenían problemas con los compañeros y, excepto uno de los chicos, eran personas bien aceptadas entre sus compañeros y tenían amigos que los apreciaban. Percibían que *la imagen que sus amigos tenían de ellos era positiva* y todos estaban convencidos de que sus compañeros y amigos tenían una buena imagen de ellos/as.
3. Tenían una *percepción positiva del ambiente de aprendizaje*. Reconocían que había profesores y profesoras que les gustaban y otros que no, y sabían explicar las cualidades de sus profesores/as preferidos/as. Valoraban la actitud de los profesores ante el aprendizaje de los alumnos, el esfuerzo que dedican a ayudar a los alumnos a aprender y también que conocieran individualmente a los estudiantes que están a su cargo.
4. A lo largo de su vida estos chicos y chicas habían encontrado personas mayores que ellos que habían llamado su atención y a los cuáles habían considerado, en algunos aspectos, *modelos a seguir*. Así, los educadores y las educadoras, los pro-

15 En Cataluña, la función de tutela o guarda ejercida por la administración en casos de niños o niñas atendidos en centros residenciales de acción educativa, dependientes del Departamento de Bienestar Social, finaliza al cumplir los jóvenes la mayoría de edad. (18 años). A partir de este momento el/la joven deberá vivir de forma autónoma e independiente. Esto hace que en muchos casos la última etapa de residencia en el centro esté marcada por múltiples tensiones relacionadas con la búsqueda de empleo, de lugar de residencia, etc.

fesores y las profesoras, una hermana mayor o el padre en algún caso, marcaron pautas que los jóvenes decidieron intentar seguir. Estas personas significativas les sugirieron formas de actuar y actitudes positivas hacia el estudio, hacia el entorno, y formas positivas de relacionarse con los demás.

En relación a los resultados obtenidos en los cursos anteriores todos *valoraron positivamente dichos resultados*. Los profesores y los educadores también los valoraban positivamente, independientemente de cuales hubieran sido las calificaciones obtenidas. Tomando en consideración las circunstancias de cada uno de estos estudiantes, *valoraban de forma positiva el esfuerzo que debían hacer para superar los requisitos de los cursos*, aunque también señalaron que los resultados, en la mayoría de los casos, podrían haber sido todavía mejores si se hubieran puesto más empeño o si no hubiesen existido circunstancias desfavorables como el hecho de tener que compaginar los estudios con un trabajo.

Analizando el complejo proceso que habían seguido estos jóvenes y desde la perspectiva de considerar el papel activo que éstos tienen en su aprendizaje, observamos que, especialmente dos de las chicas, habían sabido enfrentarse a su realidad para intentar superarla, desarrollando, en el caso de los estudios, los mecanismos para situarse ante el aprendizaje con una actitud positiva a pesar de los condicionantes familiares adversos. Ambas eran conscientes de su situación. Ante las condiciones desfavorables supieron movilizar sus propios recursos e incluso hallar en esta situación la motivación para superarla. Parece que el principal elemento que contribuyó a que las chicas movilizaran sus recursos fue el hecho *de tener un objetivo, una meta, un proyecto en su vida en el que los estudios jugaban un papel importante*.

Para alcanzar este objetivo las jóvenes pusieron en marcha su sentido de la responsabilidad. Sentirse *responsables de su aprendizaje* constituyó también un factor protector. En este sentido, deseaban superarse, estaban motivadas para el logro, sabían que debían esforzarse para tener éxito en los estudios y que el control de este esfuerzo estaba en sus manos. Habían desarrollado algunos hábitos relacionados con el estudio, aunque no quedó demasiado claro en nuestro estudio que hubieran ejercido una función como factor protector.

El *autoconcepto positivo y su autoestima* también constituyeron factores protectores. Su autoestima se reforzaba gracias al hecho de conocer lo que pensaban de ellas los profesores y los educadores. Se preocupaban por aparecer con una buena imagen, les importaba lo que los demás pensasen de ellas y esto provocaba a su vez que los demás desarrollasen sentimientos positivos hacia ellas. El deseo de ser aceptadas las aproximaba a la gente y les llevaba a establecer relaciones de amistad. Al sentirse aceptadas se reforzaba su autoestima. Las *habilidades para relacionarse e interactuar con los demás* podían haber ejercido también la función de factores protectores.

Ser aceptados y aceptadas en la escuela y el instituto hizo que los jóvenes se sintieran bien en este ambiente. Saber que tanto profesores/a como educadores/as y compañeros/as tenían conceptos y actitudes positivas hacia ellos les ayudó a desarrollar actitudes positivas hacia la escuela y hacia el aprendizaje.

4.6. Discusión

El objetivo del estudio de casos y de la comparación entre ellos era hallar elementos comunes que nos pudieran sugerir nuevas hipótesis para la investigación y nuevas posibilidades para la acción educativa escolar, especialmente orientada a la prevención del fracaso escolar. Resumiendo todo lo dicho, el estudio de casos nos aporta información que indica como posibles factores protectores del riesgo de fracaso escolar en los casos estudiados los que siguen:

- Ser *consciente de la propia situación compleja y desfavorable*, y de las repercusiones que esto puede tener en todos los aspectos de su vida, incluidos los estudios.
- Tener *un objetivo, una finalidad o un proyecto sobre sí mismo a medio o largo plazo*, que constituye la motivación principal, y en el cual los estudios son percibidos como un mecanismo útil y necesario para alcanzarlo.
- Tener un *autoconcepto positivo y una autoestima alta*.
- Sentirse *autoresponsables de su aprendizaje*.
- Tener *habilidades para interrelacionarse con los demás*.
- Identificar en algunas personas próximas *modelos a seguir*, como referentes.
- Tener *actitudes positivas hacia la escuela y, especialmente, hacia el aprendizaje*.
- En menor proporción, el haber desarrollado ciertos *hábitos de estudio*, especialmente de organización y de constancia.

De estos factores, los dos primeros no habían sido incluidos en el modelo de riesgo como posibles variables a tener en cuenta. Según se desprende de este estudio, estos dos factores constituirían las condiciones previas para lograr desarrollar unas actitudes positivas hacia la escuela, un nivel importante de responsabilidad hacia el propio aprendizaje, y ciertos hábitos para el estudio. Tampoco el hecho de percibir determinadas personas como modelos había sido previsto como factor protector y, en cambio, del análisis de los casos se desprende que los chicos y las chicas se fijan en los roles y las conductas de otras personas significativas para ellos y se proponen seguirlas.

Una buena parte de estos resultados concuerdan con los hallados en otras investigaciones sobre *resilience*. Tener una autoestima positiva y un buen concepto de sí mismos es uno de los aspectos coincidentes con la mayoría de las investigaciones revisadas (Garmezy, 1983¹⁶; Connell, Spencer *et al*, 1994; Werner y Smith, 1994¹⁷; Clark, 1995). Las actitudes positivas hacia la escuela son destacadas como factor protector por Werner y Smith (1994), los cuales señalan como factor protector el hecho de considerar la escuela como una experiencia valiosa, y por Clark (1995) para quien la motivación y las aspiraciones positivas son algunos de los aspectos que ayudan a compensar el riesgo.

¹⁶ Citado por Winfield, 1995, op. cit.

¹⁷ Citado por Burger, 1994, op. cit.

El sentido de responsabilidad personal también es señalada por algunos autores como un factor vinculado a la *resilience* aunque no únicamente relacionado con la escuela y el aprendizaje. De este modo, los autores señalan el *sentido de poder personal* y el *locus of control interno* (Garmezy, 1983), el sentido de control sobre sus éxitos y sus fracasos en la escuela (Connell, Spencer *et al*, 1994), la habilidad para actuar independientemente y ejercer control sobre el propio entorno (Berliner y Benard, 1995), y la responsabilidad personal en sentido genérico (Clark, 1995). La mayoría de los autores mencionados coinciden además en que las personas «resilientes» poseen habilidades para relacionarse con los demás y la *capacidad para establecer relaciones positivas*. Tanto Garmezy (1983), como Miller (1995), Werner y Smith (1989) y Clark (1995) destacan el hecho de que las personas «resilientes» han tenido un adulto significativo en sus vidas y un amigo especial. Esto coincide con los resultados de nuestra investigación. En el caso de los jóvenes de nuestro estudio, además de establecer relaciones positivas y significativas con adultos parece que los jóvenes ven en algunos de estos adultos significativos modelos a seguir. Esto resulta especialmente importante en el caso de los jóvenes de nuestro estudio los cuales tienen un escaso vínculo con su familia biológica y cuyos referentes adultos son los educadores y educadoras y los profesores y profesoras con los que han establecido un vínculo.

En cuanto a ser conscientes de su propia situación desfavorable éste no es un aspecto que haya sido considerado de forma general. En la revisión realizada únicamente Miller (1995) que realizó una investigación sobre la «resiliencia» con jóvenes universitarios con dificultades de aprendizaje destacó como un factor a tener en cuenta el reconocimiento de la dificultad de aprendizaje por parte de los jóvenes, y Berliner y Benard (1995) mencionan el tener sentido de la propia identidad como aspecto a considerar, que podría incluir la consciencia de la propia situación. Sin embargo, creemos que las investigaciones revisadas no reflejan la importancia que puede tener este aspecto como factor protector en casos de alto riesgo. Una imagen realista de sí mismos y de la situación que viven, de sus límites, de sus potencialidades y de sus condicionantes, favorecerá la articulación de mecanismos adecuados para hacer frente a esta situación. El conocimiento de la situación puede ser un punto de partida necesario para que el individuo articule respuestas orientadas a superar las dificultades.

El tener un objetivo a medio o largo plazo que orienta y motiva a estos muchachos y muchachas y que ha aparecido como resultado de la investigación estaría a nuestro entender relacionado con el sentido de futuro o habilidad para imaginarse un futuro propio, «ser optimista» (Berliner y Benard, 1995) y a la visión positiva del futuro personal citado por Clark (1995). El proyecto o meta del que nosotros hablamos no sería otra cosa que tener una visión de futuro optimista, puesto que de lo contrario no constituiría una motivación para estos jóvenes.

Los hábitos de estudio solamente son mencionados por Werner y Smith (1989) que citan la concentración y capacidad de atención como factor asociado a la *resilience*. Sin embargo, la mayoría de las investigaciones revisadas señalan las habilidades para resolver problemas, tomar decisiones, buscar ayuda en los demás cuando la tarea lo requiere, como características de las personas «resilientes» (Garmezy, 1983; Werner y Smith, 1989; Berliner y Benard, 1995; Clark, 1995). En nuestro caso no hemos podido

determinar en los jóvenes de nuestro estudio que este tipo de habilidades haya jugado un papel relevante como factor protector ni que sea una característica significativa, sin embargo esto puede deberse a las limitaciones de nuestra investigación y de los instrumentos utilizados para hallar indicios de la existencia de este tipo de habilidades en los jóvenes.

En síntesis, los resultados concuerdan con los obtenidos en otras investigaciones sobre las características de las personas «resilientes». A pesar de que los resultados de nuestra investigación, en coherencia con su punto de partida, indican algunas características individuales como posibles factores protectores, no debemos olvidar que las respuestas individuales no serán suficientes para prevenir el riesgo de fracaso escolar y que será necesario articular entornos que ayuden a potenciar estas respuestas en los individuos. La delimitación de estas variables como factores protectores del fracaso escolar en casos de alto riesgo nos permitirá apuntar algunas líneas de acción educativa destinadas a la prevención del fracaso escolar, que en su mayor parte consistirán en crear los entornos adecuados para potenciar en los estudiantes factores protectores.

5. CONCLUSIONES

Desde nuestro punto de vista, este enfoque de investigación que hemos presentado resulta interesante no solamente por su capacidad para aportar conocimientos significativos para diseñar estrategias preventivas que contribuyan a disminuir el riesgo de fracaso escolar, sino también por una cuestión metodológica como es el hecho de que supone un análisis del potencial del «caso atípico» como base para la investigación educativa.

La investigación sobre factores protectores, es decir, sobre factores que han contribuido a la «resiliencia» de individuos que se encuentran en situación de riesgo, nos aporta datos sobre las características que presentan estas personas y las características de su entorno. Aunque surgen resultados interesantes no debemos perder de vista que lo serán realmente en la medida en que, a partir de su conocimiento, podamos articular intervenciones educativas para prevenir el fracaso escolar, tanto en niños y niñas en situación de riesgo como en la población escolar en general. Las intervenciones educativas deberán ir destinadas a promover los factores protectores, tanto personales como del entorno.

En algunos casos, se ha intentado trabajar directamente aspectos como la autoestima de los estudiantes y su autoconcepto utilizando programas específicos. Sin embargo, como señala Winfield (1996), este tipo de intervenciones tiene efectos a corto plazo y difícilmente se producirán cambios estables en los niños. Para esta autora las estrategias de intervención deben enfocarse de una manera más amplia, potenciando en los estudiantes procesos protectores en momentos críticos de su vida. La «resiliencia» es algo que debe irse potenciando a lo largo de la vida de las personas.

Algunos autores coinciden en la necesidad de desarrollar estrategias que ayuden al autoconocimiento de la propia situación y la autocomprensión de la misma como una

base para tomar las riendas de la propia vida (Miller, 1995), que ayuden a la estimación realista de las propias capacidades y a aumentar las creencias en las propias habilidades y capacidades (Connell, Spencer y Aber, 1994; Miller, 1995). Connell *et al* (1994) consideran como mecanismo facilitador de la «resiliencia» la promoción de la implicación de la familia con la escuela. Otros autores como Berliner y Benard (1995) y Clark (1995) se centran en prácticas educativas que promuevan la *resilience* desde la misma escuela.

No debemos perder de vista que los factores de riesgo y los factores protectores están vinculados al contexto o a la comunidad en la que el individuo se desarrolla. Por lo tanto, hay que contemplar la ayuda a la infancia en situación de riesgo desde una perspectiva más amplia que la escolar. Así, aunque podamos coincidir con otras investigaciones en cuanto a características de las personas «resilientes», y estemos de acuerdo en que es necesario abordar la intervención desde un enfoque global, es necesario pensar en estrategias de intervención contextualizadas. Esto significa que hay que estudiar en profundidad la comunidad en la que está involucrado el individuo para poder determinar estrategias de intervención educativa y social.

En el contexto en el que hemos situado nuestra investigación existen servicios a la comunidad como los servicios sociales de atención primaria, los Equipos de Atención a la Infancia en situación de riesgo que trabajan desde el Departamento de Bienestar Social, las escuelas con sus Equipos de Asesoramiento Pedagógico. La coordinación entre todos estos servicios y de los profesionales que los integran (educadores/as sociales, pedagogos/as, asistentes sociales, psicólogos/as, trabajadores/as familiares, profesores/a) debería ser el punto de partida para promover en los niños y las niñas los recursos que les permitieran hacer frente a las situaciones de riesgo. En nuestro entorno existen ya experiencias de coordinación pero todavía falta mucho camino por recorrer.

Partiendo de nuestra investigación podemos decir que la intervención de los profesionales de la educación en la *prevención del fracaso escolar en niños y niñas en situación de riesgo* debería considerar, al menos, los siguientes objetivos:

- *Ayudar al niño a tomar consciencia de su situación.* A grandes rasgos, implica que los profesionales deben ser realistas con los niños tanto en lo referente a la situación actual como a su futuro. Es importante que las intervenciones educativas desarrollen en los niños y niñas recursos para hacer frente a su realidad.
- *Ayudar al niño a desarrollar un proyecto vital.* Este es uno de los recursos personales que el niño puede desarrollar con la ayuda de profesores y educadores sociales. Evidentemente, no podemos pretender que un niño pequeño tenga un proyecto vital. Pero sí que podemos ayudarle a que aprenda a proponer objetivos de sus acciones a corto plazo y, posteriormente a medio y a largo plazo. Este objetivo debería priorizarse en algunos momentos críticos de la vida de la persona, como la transición a la vida activa e independiente.
- *Ayudar al niño a desarrollar el sentido de la responsabilidad sobre su aprendizaje y, en general sobre su propia vida.* Los profesionales deben contribuir a que los niños perciban que tienen control sobre su aprendizaje y sobre su vida en general, que

aprendan a atribuir la causa de sus éxitos y de sus fracasos en el aprendizaje a variables controlables como el esfuerzo y no a la falta de capacidades o a la suerte. Si sienten que pueden influir en su propia vida seguramente serán más capaces para hacer frente a circunstancias adversas, serán más responsables y estarán más motivados en todas las dimensiones de su vida: escolar, laboral y afectiva.

- *Ayudar al niño a desarrollar un autoconcepto y una autoestima positivos.* Hay que ayudarlos a valorarse a sí mismos, a que sepan reconocer sus cualidades y sus capacidades y habilidades.
- *Potenciar las actitudes positivas hacia la escuela y el aprendizaje.* Estas actitudes tendrán más posibilidades de desarrollarse si se ofrecen al niño estrategias y habilidades para llevar a cabo sus aprendizajes. Debemos tener en cuenta, además, que las actitudes de los profesionales de la educación repercutirán en la imagen que el niño se forma de la escuela y del aprendizaje.
- *Ayudar al niño a desarrollar habilidades de interrelación social.* Los profesionales tendrían que incidir en el desarrollo de estas habilidades por la influencia que tienen en el concepto y las expectativas que los demás se forman de ellos.
- *Potenciar el desarrollo de hábitos de estudio.* Los profesionales de la educación deben ayudar a estos niños y niñas a desarrollar hábitos de constancia, organización y atención en el trabajo y el aprendizaje escolar, pero también en otras dimensiones de su vida.

En resumen, desde el enfoque que hemos presentado la investigación no debe centrarse únicamente en buscar características de las personas «resilientes» sino en buscar qué estrategias y qué intervenciones ayudan a disminuir el riesgo y a promover la «resiliencia» de las personas. Así, nuestra investigación sería un primer paso para conocer qué aspectos deben potenciarse mediante intervenciones educativas adecuadas dentro del marco escolar y en coordinación con los servicios de la comunidad. Teniendo en cuenta que para definir ciertas variables como factores de riesgo deben tener posibilidades de modificación mediante la intervención educativa preventiva, es necesario que la investigación ayude a comprobar si la intervención sobre determinados factores de riesgo o sobre factores protectores puede contribuir a disminuir el riesgo de fracaso escolar.

En el *aspecto metodológico* la investigación sobre «resilience» pone énfasis en el estudio de casos que podríamos denominar «atípicos». Con frecuencia la investigación educativa se ha preocupado por obtener informaciones sobre lo general y común. Ello ha aportado conocimientos valiosos pero también puede aportarlos el estudio de casos singulares y atípicos. La dificultad de las investigaciones sobre riesgo y *resilience* radica en que para poder abordar la mayoría de aspectos que puedan resultar relevantes es necesario un estudio en profundidad, a largo plazo y contextualizado de los casos. En este sentido, las investigaciones longitudinales y de estudio de casos podrían aportar información significativa sobre variables que ejercen una función de factores protectores.

6. BIBLIOGRAFÍA

- Ball, S. (1988). *La motivación educativa. Actitudes. Intereses. Rendimiento. Control*. Madrid: Narcea.
- Berliner, B.; Benard, B. (1995). *More than a Message of Hope: A District-Level Policymaker's Guide to Understanding Resiliency*. Western Regional Center for Drug-Free Schools and Communities, Portland. ERIC Document ED 387 946.
- Bardin, L. (1986). *Análisis de contenido*. Madrid: Akal Universitaria.
- Burguer, J.V. (1994). Keys to Survival: Highlights in Resilience Research. *Journal of Emotional and Behavioral Problems*, 3 (2), 6-10.
- Casas Aznar, F. (1994). Prevenció i qualitat de vida dels infants. *Revista d'informació i investigació socials*, 0, 40-47.
- Clark, P. (1995). *Risk and Resiliency in Adolescence: The Current Status of Research on Gender Differences*. Ohio State University, Columbus Dept. of Home Economics Education. ERIC Document: ED 387 714.
- Connell, J.P.; Spencer, M.B.; Aber, J.L. (1994). Educational Risk and Resilience in African-American Youth: Context, Self, Action and Outcomes in School. *Child Development*, 65 (2), 493-506.
- Forns Santacana, M. (1989). L'avaluació de l'aprenentatge. Conferència ofrecida en Tarragona en el *Seminari sobre Aprenentatge*, organitzat per el Departament d'Educació i Psicologia, Àrea de Psicologia Bàsica, y l'Associació de Psicòlegs de Tarragona. Tarragona, 6 i 7 d'abril de 1989.
- Frey, D.K.; Jesse Carlock, J. (1989). *Enhancing Self Esteem, Accelerated Development* INC. Publishers. Muncie.
- Goetz, J.Pl.; LeCompte, M.D. (1988). *Etnografía y diseño cualitativo en investigación educativa*. Madrid: Morata.
- González, M.C.; Tourón, J.: *Autoconcepto y rendimiento escolar*. Pamplona: Eunsa.
- Gordon, K.A.; Ingersoll, G.M.; Orr, D.P. (1995). Profile of Behaviorally Resilient Adolescents: Confirmation and Extension. Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, april, 18-22, 1995). ERIC Document. ED 385 810.
- Guetzloe, E. (1994). Risk, Resilience, and Protection. *Journal of Emotional and Behavioral Problems*, 3 (2), 2-5.
- Manciaux, M. (1994). De la vulnerabilitat a la resiliència, i dels conceptes a l'acció. *Revista d'informació i investigació socials*, 0, 48-56.
- Miller, M. (1995). *Sources of Resilience Outcomes*. Paper presented at the International Convention of the Council for Exceptional Children. (73rd, Indianapolis, A april 5-9 1995) ERIC Document. ED 381 949.
- Pianta, R.C. (1990). Widening the Debate on Educational Reform: Prevention as a Viable Alternative. *Exceptional Children*, 56 (4), 306-313.
- Rodríguez Espinar, S.; Álvarez, M.; Echeverría, B.; Marín, M.A. (1993). *Teoría y práctica de la Orientación Educativa*, Barcelona: PPU.
- Wang, M.C.; Peverly, S.T. (1987). The Role of the Learner: An Individual Difference Variable in School Learning and Functioning. Wang, M.C.; Reynolds, M.C.; Wal-

- berg, H.J. (Eds.). *Handbook of Special Education. Research and Practice. Vol. 1: Learning Characteristics and Adaptive Education.* (pp. 59-92.) Oxford: Pergamon Press.
- Winfield, L.F. (1996). Developing Resilience in Urban Youth. *ncrel Monograph.* <http://ousd.k12.ca.us/netday/links/Partnerships/le0win.htm>. (17 junio 1996).
- Yin, R.K. (1990). *Case Study Research. Design and Methods.* London: Sage Publications.
- Zimmerman, B.J.; Martínez-Pons, M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Reserach Journal*, 23, 614-628.
- Zimmerman, B.J.; Martínez-Pons, M. (1988). Construct Validation of a Strategy Model of Student Self-Regulated Learning. *Journal of Educational Psychology.* 80 (3), 284-290.

LA ESCALA KEYMATH-R: ESTUDIO DE UNA PRUEBA DIAGNÓSTICA CON REFERENCIA CURRICULAR

Ricard Marí Mollà
Universitat de València¹

RESUMEN

Las nuevas tendencias de la investigación exigen a los instrumentos de evaluación y diagnóstico una adecuación a las nuevas orientaciones de la enseñanza y una fundamentación en los propios procesos de aprendizaje. En el caso del aprendizaje de las matemáticas y desde un enfoque criterial, se presenta una revisión de la escala KeyMath-R, una prueba referida a los contenidos o dominios, que ha mostrado una organización interna correspondiente al currículum, con unos niveles de bondad métrica muy aceptables, considerándose un instrumento diagnóstico de gran utilidad.

ABSTRACT

The new investigation trends require to the evaluation and diagnostic tools a learning new trainings adaptation and foundation in the own learning process. For the mathematics achievement evaluation and with criterion-referenced, is made a KeyMath-R scale review. This is a domain-referenced inventory, with an internal organization corresponding to the curriculum. The reliability and validity coefficient levels show enough acceptability, which support its use.

Desde hace más de una década el nombre de **KeyMath-R** es un sinónimo de investigación educativa y está presente en estudios sobre tests de rendimiento y en los listados de pruebas educativas. Es una escala para evaluar las habilidades y capacidades

¹ Ricard Marí Mollà es Profesor Titular de la Universitat de València, adscrito al Departamento MIDE. E-meil: Ricard.Mari@uv.es

matemáticas y particularmente útil para el diagnóstico de los déficits académicos en sujetos con dificultades de aprendizaje y, además, por su adecuación curricular permite diseñar programas de intervención educativa.

Su éxito se deriva de su alcance comprensivo, de su contenido motivador y colorista, y de la facilidad de su administración, puntuación y aplicación a la práctica instruccional. El elemento principal lo integra el *KeyMath Revised: A Diagnostic Inventory of Essential Mathematics* (Connolly, 1988).

El KeyMath-R forma parte de un conjunto de varios elementos: un cuestionario diagnóstico en sus formas A y B, un suplemento métrico a la prueba, un programa de desarrollo del primer nivel, un programa de rehabilitación y un manual de actividades manipulativas y ejercicios (Connolly, 1985, 1988 y 1988a).

El **KeyMath-R** es una segunda construcción de una prueba de evaluación de logros de aprendizaje en matemáticas elementales realizada por A.J. Connolly, el *Key-Math*, en este caso mejorada notablemente en sus aspectos de amplitud de aplicación, de adaptación curricular y normalización, para que pueda ser utilizado con todo tipo de sujetos, normales o con dificultades de aprendizaje y en diferentes situaciones de aprendizaje, tanto para la evaluación en el aula normal como en el ámbito del diagnóstico de las dificultades de aprendizaje y para todos los niveles de la escuela primaria.

Está considerado como un «Inventario Diagnóstico de Matemáticas Básicas». Una escala de aplicación individual que evalúa las habilidades matemáticas correspondientes a los diferentes niveles de la educación primaria y que consta de 13 Subpruebas clasificadas en 3 Subescalas: Conceptos Básicos, Operaciones (procesos de cálculo) y Aplicaciones. Es una prueba con referencia a los contenidos o dominios y validado para la interpretación normativa. Según Bachor (1989-90) en diferentes estudios sobre el KM-R se le ha considerado por los autores como test normativo, por las tablas normativas del manual, como test referido al dominio e, incluso, como test referido al criterio (Tinney, 1975; Price, 1984).

El interés de la prueba, aparte de su validez demostrada durante muchos años en los E.E.U.U. en el diagnóstico de las dificultades de aprendizaje de las matemáticas, tanto en su construcción original como en la revisión que estamos analizando, radica en que identifica en cada momento el contenido esencial (conceptos y destrezas) del trabajo del alumno, organizado en un marco conceptual que facilita la realización del diagnóstico descriptivo y la planificación de la futura instrucción.

Se trata de un instrumento de administración individual, que entre 35 y 50 minutos puede ofrecer una evaluación comprensiva del conocimiento y aplicación de conceptos o habilidades matemáticas importantes de un alumno. Si bien la intención del KM original estaba clara, como se ha dicho, el propósito del KM-R no ha sido definido por el autor, no se sabe si su diseño está pensado para los alumnos especiales o los normales. No obstante, siguiendo a Bachor (1989-90), según la selección de sujetos efectuada en el proceso de estratificación de las variables siguiendo el diseño de validación, se considera adecuado para ser utilizado tanto con sujetos considerados normales, como con sujetos afectados por variables raciales o étnicas que supongan deficiencias culturales y/o lingüísticas.

ELEMENTOS DE LA ESCALA KEYMATH-R

Alcanza un rango amplio de edades: desde Preescolar, hasta el noveno grado. Dispone de dos formas alternas, A y B, equiparadas estadísticamente y por contenido. Su contenido se basa en un sistema comprensivo referido al dominio que identifica jerarquías de conceptos y habilidades a través de trece subpruebas de contenido. Cada una de las 13 subpruebas del sistema se mide por separado. Las subpruebas, por su parte, se agrupan en áreas de contenido matemático y evalúan dicho contenido en las tres subescalas ya mencionadas de: Conceptos Básicos, Operaciones y Aplicaciones (Ver Tabla 1).

Tabla 1
ESTRUCTURA DE CONSTRUCCIÓN DE LA ESCALA KM-R

Conceptos Básicos	Operaciones	Aplicaciones
Numeración	Suma	Medida
Número Racional	Resta	Tiempo y Dinero
Geometría	Multiplicación	Estimación
	División	Interpretación datos
	Cálculo mental	Resolución problemas

Cada subprueba contiene tres o cuatro dominios con una importancia instruccional similar. Cada uno de estos dominios se representa por un conjunto de seis ítems. Con esta representación el KM-R puede ofrecer datos fiables por debajo del nivel de subprueba, los cuales, según el constructor, facilitan el análisis y la traslación de los resultados de la prueba a alternativas instruccionales.

Ofrece información diagnóstica a cuatro niveles: Total Escala, Subescalas, Subpruebas y Dominios. Cada nivel ofrece una información específica sobre la ejecución del alumno al identificar los puntos fuertes y débiles, es decir, lo que sabe y lo que no sabe el alumno y, de esta forma, se establece una base para confeccionar un programa instruccional adecuado:

Hay que añadir, que, según las intenciones del autor, las interpretaciones de los resultados de la prueba hechas en los niveles ya mencionados, se pueden aumentar con un análisis de la ejecución del estudiante sobre ítems específicos, los cuales aparecen ordenados de menor a mayor grado de dificultad en cada uno de los dominios. Para esta finalidad, el autor proporciona, además de los ítems, un listado de los objetivos de los mismos.

CONTENIDO DE LA ESCALA KM-R

Teniendo en cuenta la exposición anterior y la definición que del mismo realiza A.J. Connolly en el manual, el KeyMath-R es una prueba referida al contenido y con provisiones para una interpretación normativa. A través de su desarrollo, el énfasis se sitúa en la identificación del contenido esencial (conceptos y habilidades) y en la organización de dicho contenido en un sistema que facilite la evaluación diagnóstica y la planificación de la futura instrucción.

En su construcción, Connolly tuvo que resolver un importante problema derivado de la búsqueda del contenido matemático esencial. A partir de la revisión de los programas matemáticos básicos se encontró con que no existía un modelo consensuado para la organización y delineación del contenido matemático en la enseñanza elemental y media. Esta dificultad se complicaba todavía más por los problemas que se presentaban al intentar unir los esfuerzos en la evaluación y rehabilitación con el desarrollo instruccional.

Así pues, la construcción del sistema de contenido la empezó Connolly con una amplia revisión de los currícula matemáticos existentes, de las prioridades instruccionales actuales y de las tendencias futuras. Este esfuerzo incluyó una revisión del alcance y secuencia de algunos libros de texto básicos de la enseñanza elemental y las recomendaciones curriculares establecidas en *An Agenda for Action* del National Council of Teachers of Mathematics (1980).

Posteriormente este problema se ha resuelto de alguna forma al imponerse las recomendaciones del NCTM en el área de las matemáticas elementales y configurarse como un estándar para los educadores (NCTM, 1988 y 1989; Tej Pandey, 1994).

La revisión de la escala KeyMath-R realizada por nosotros (Marí, 1996) incluye sendos trabajos de Validación de Contenido desde la perspectiva del Diseño Curricular tanto del estándar del NCTM como de las prescripciones para la implantación de la LOGSE, que completan y mejoran el proceso de estudio del contenido y la Validez de Constructo de la adaptación original de la escala.

ALCANCE Y SECUENCIA

El sistema de la escala KM-R comprende, pues, tres *Subescalas*: Conceptos Básicos, Operaciones y Aplicaciones, a las que se les concede la misma importancia instruccional. A su vez, las tres *Subescalas* anteriores se divide en trece *Subpruebas*, seleccionadas y desarrolladas con una importancia equivalente, y cada una de ellas, a su vez, en tres o cuatro *Dominios*, que de nuevo mantienen el principio de equilibrio en la consideración de la importancia entre los diferentes dominios.

La especificación del contenido cuidadosamente articulada y equilibrada, descrita en la Tabla siguiente (ver Tabla 2), representa la aplicación desde los currícula en la que el valor instruccional varía marcadamente a lo largo de los subpruebas o a lo largo de los dominios.

Tabla 2

ESPECIFICACIÓN DEL CONTENIDO DE LA ESCALA KEYMATH-R: SUBESCALAS, SUBPRUEBAS Y DOMINIOS

Subescala:	CONCEPTOS BÁSICOS	OPERACIONES	APLICACIONES
Subprueba: Dominios:	<p>Numeración Números 0-9 Números 0-99 Números 0-999 Números multidígitos y tópicos de numeración avanzados</p> <p>Números Racionales Fracciones Decimales Percentiles</p> <p>Geometría Relaciones espaciales y atributivas Formas bidimensionales y sus relaciones Geometría coordinada y transformacional</p> <p>Formas tridimensionales y sus relaciones</p>	<p>Suma Modelos y hechos básicos Algoritmos para sumar números enteros Suma de números racionales</p> <p>Resta Modelos y hechos básicos Algoritmos para restar números enteros Resta de números racionales.</p> <p>Multiplicación Modelos y hechos básicos Algoritmos para multiplicar números enteros Multiplicación de números racionales comerciales</p> <p>División Modelos y hechos básicos Algoritmos para dividir números enteros División de números racionales</p> <p>Cálculo mental Cálculos encadenados Números enteros Números racionales</p>	<p>Medida Comparaciones Uso de unidades no estándar Uso de unidades estándar: longitud, y área Uso de unidades estándar: peso y capacidad</p> <p>Tiempo y dinero Identificación del paso del tiempo Uso de relojes y de las horas Uso monetario de un dólar Uso monetario de cien dólares y transacciones</p> <p>Estimación Números racionales y enteros Medidas Cálculo</p> <p>Interpretación de Datos Cuadros y tablas Gráficos Probabilidad y estadísticos</p> <p>Solución de problemas Resolver problemas rutinarios Comprensión de problemas no rutinarios Resolver problemas no rutinarios</p>

El KM-R es, pues, un instrumento de evaluación basado en un sistema específico, cada apartado de conocimiento de las tres subescalas del KM-R se mide por una subprueba separada. Cada dominio se mide por un conjunto de seis ítems, a través de los que se nos da el nivel alcanzado por el estudiante en la realización de la prueba. El rigor aplicado a la selección y determinación del contenido, expuesto anteriormente, se empleó, también, en la formulación y desarrollo de los ítems. Cada ítem fue sometido a un estudio, consulta de expertos y demostraciones apropiadas de sus características psicométricas. Los ítems que superaron estas pruebas son los que se presentan en las diferentes subpruebas del KM-R distribuidos secuencialmente, no por los dominios que representa y pretende evaluar cada uno de ellos, sino siguiendo un orden de dificultad previamente establecido.

ESTUDIO MÉTRICO DIFERENCIAL DE LA ESCALA KEYMATH-R

La necesidad de disponer de instrumentos diagnósticos que atiendan a la totalidad tanto de los contenidos como de los elementos procesuales del aprendizaje de las matemáticas, así como poder disponer de instrumentos diagnósticos que proporcionen elementos para la intervención pedagógica y superen las deficiencias de los tests normativos al uso, ha aconsejado el estudio de algunas pruebas orientadas a la evaluación diagnóstica. El KeyMath-R es uno de dichos instrumentos diagnósticos. Tal como se ha señalado, por su referencia a los contenidos o dominios, las descripciones de los campos de conducta a evaluar, la adecuación curricular de los dominios propuestos en línea con las novedades curriculares, su impacto de uso en el campo diagnóstico, los estudios sobre criterios de bondad y el volumen de investigación desarrollado sobre el mismo, lo hacen un instrumento valioso, por lo que se ha decidido un análisis en profundidad desde nuestro ámbito cultural.

En primer lugar, dado que la escala KM-R sólo muestra una referencia curricular y para dotarla de una mayor riqueza diagnóstica, se ha realizado una revisión teórica de algunas de las conclusiones aportadas por la investigación desde el enfoque del Procesamiento de la Información y referidas a los procesos de la enseñanza y aprendizaje de las matemáticas en la escuela primaria. Con ello dotamos de una fundamentación teórica a la escala *KeyMath-Revised*, la cual otorga a la prueba una mayor riqueza de interpretación diagnóstica y permiten enriquecer los programas de intervención con las conclusiones de la propia psicología cognitiva.

En segundo lugar, se realiza un estudio empírico y de adaptación del cuestionario basado en una estructura de análisis clásico de construcción de pruebas, abordando aquellos componentes que se han considerado válidos para la depuración de un instrumento de medida: descripción de la muestra utilizada y su rendimiento, fiabilidad, validez y tipificación. Se ha tomado como referencia el modelo clásico, aunque algunos autores actualmente hacen hincapié en otros elementos como el TRI (Teoría de Respuesta al Ítem). Se ha insistido más en el ajuste del instrumento en aspectos de validez, primando elementos de análisis de contenido sobre los de depuración métrica (ej.: modelo Rasch de fiabilidad), puesto que se ha considerado de

mayor interés interpretativo dadas las características de referencia criterial de la prueba.

La **muestra** de sujetos para la adaptación del cuestionario KeyMath-R está constituida por un total de 1.238 sujetos de la Comunidad Valenciana, seleccionados al azar, de los que el 50,2% son mujeres y el 49,8% varones con un rango de edad comprendido entre los 5 y los 15 años. Todos ellos son estudiantes de los cursos comprendidos entre 2º de Educación Preescolar, EGB y 1º de BUP-1º FP.

La muestra seleccionada es representativa de la población escolar comprendida entre los niveles de Preescolar, de la EGB y 1º BUP-FP, a nivel de la Comunidad Valenciana, y se realizó sobre el informe de la población escolar de la Conselleria d'Educació i Ciència de la Comunitat Valenciana correspondiente al curso 1989-90, con el fin de asegurar una proporcionalidad estricta a la población escolar mencionada.

Para la configuración de la muestra se han tenido en cuenta diversas variables de estratificación como: curso o nivel de estudios, sexo, tipo de colegio, clase social, lengua, etc., de forma que se mantenga la distribución interna de cada grupo de edad, conservando las mismas proporciones en la muestra que en la población.

El estudio empírico incluye, preceptivamente, un **análisis de ítems**. Este análisis se ha realizado desde tres perspectivas. En primer lugar, una revisión empírica de los niveles de acierto mediante el análisis del índice de dificultad (π) a lo largo de cada Subprueba para cada uno de los cursos y el total. En segundo lugar, la comprobación de si el ordenamiento de los ítems en cada Subprueba responde a una gradación de dificultad. Y, en tercer lugar, una síntesis comparativa que nos permita encontrar la adecuación entre los niveles de dificultad observados en la revisión empírica y los niveles de logro teórico, que determina el perfil teórico previsto, derivados de la revisión lógica realizada por un grupo de profesores de matemáticas seleccionados, indicando el ciclo o nivel en que cada ítem debe estar dominado según el Diseño Curricular Base.

Se comprueba que los ítems y los dominios de la Escala están graduados con un nivel de dificultad progresiva y que los niveles de rendimiento de los estudiantes coinciden con el perfil teórico de logro.

Estos análisis permiten realizar interpretaciones diagnósticas a nivel de ítem y de dominios, pudiéndose analizar los logros de los estudiantes independientemente de la interpretación de la puntuación global de la Subprueba.

Para la determinación de la consistencia de las puntuaciones obtenidas en el pase de las Subpruebas de la Escala KeyMath-R, se ha realizado un estudio de **fiabilidad**, mediante los coeficientes α de consistencia interna (Cronbach KR-20, teniendo en cuenta que son dicotómicos) y el procedimiento de las dos mitades (Spearman-Brown), para cada Subprueba y para cada Dominio, en cada uno de los niveles educativos estudiados. Los resultados obtenidos pueden considerarse buenos, con unos coeficientes mayores que, en general, a .80 y superiores a los obtenidos por Connolly en la Escala original, comprobándose la correspondencia de los coeficientes con el nivel de dificultad y con los niveles de instrucción-logro. Este hecho se considera una manifestación lógica de la adecuación de la prueba al Diseño Curricular.

Los diferentes coeficientes de fiabilidad referidos a las Subescalas, Subpruebas y Dominios, teniendo en cuenta los efectos que la variabilidad produce sobre los mis-

mos, presentan unas secuencias evolutivas que interpretamos relacionadas con la instrucción de los estudiantes.

Como ejemplos presentamos los coeficientes de fiabilidad referidos a las Subpruebas y Áreas (ver Tablas 3 y 4) en los que se observa la secuencia mostrada por los coeficientes destacando los cursos inferior (Preescolar) y superior (1º BUP), con unos niveles sin instrucción o con instrucción completa, que muestran, por ello, los coeficientes más bajos. Todo ello se considera la manifestación lógica de la adecuación de la prueba al Diseño Curricular.

Los análisis sobre la validez necesarios para la adaptación de la escala KeyMath-R, han versado sobre sendos estudios de **Validez de Constructo** y **Validez de Criterio**.

Dadas las características del KeyMath-R como prueba de rendimiento y referida al currículum, los análisis de la Validez de Constructo se han interpretado desde la perspectiva del Diseño Curricular y en la búsqueda de Evidencias de Validez de Contenido y de Validez Curricular necesarias para una interpretación de tipo criterial, mediante unos procedimientos más potentes que los utilizados por Connolly en la edición original. Así, se ha realizado una revisión de trabajos sobre validación de contenido, cuyo criterio de validación se ha centrado en la correspondencia de los dominios de la prueba con el currículum. En primer lugar, y siguiendo la estrategia de Romberg

Tabla 3

*COEFICIENTES DE FIABILIDAD S-B DE LAS SUBPRUEBAS DEL KM-R
CORRESPONDIENTES A LOS DIFERENTES CURSOS DE NUESTRA MUESTRA*

	Coeficientes de Fiabilidad Spearman-Brown										
	Preesc.	1º	2º	3º	4º	5º	6º	7º	8º	1ºBUP	1ºFP
Subprueba 1	.83	.91	.91	.89	.89	.92	.86	.80	.86	.32	.94
Subprueba 2	.74	.75	.90	.84	.93	.89	.91	.90	.92	.82	.92
Subprueba 3	.81	.90	.91	.89	.87	.87	.86	.81	.89	.73	.91
Subprueba 4	.87	.92	.87	.85	.96	.74	.76	.77	.73	.43	.68
Subprueba 5	.53	.88	.89	.85	.94	.82	.76	.82	.78	.51	.72
Subprueba 6	.56	.73	.91	.80	.73	.79	.76	.80	.77	.66	.73
Subprueba 7	.59	.69	.96	.89	.81	.92	.79	.89	.85	.82	.93
Subprueba 8	—	.83	.88	.87	.92	.94	.88	.90	.88	.75	.92
Subprueba 9	.86	.91	.84	.85	.95	.90	.82	.84	.86	.64	.85
Subprueba 10	.16	.91	.92	.91	.97	.91	.90	.84	.86	.64	.85
Subprueba 11	.52	.69	.77	.72	.92	.83	.89	.84	.79	.80	.70
Subprueba 12	.76	.84	.85	.78	.93	.90	.87	.83	.86	.67	.70
Subprueba 13	.74	.82	.79	.79	.86	.88	.88	.86	.86	.67	.78

Tabla 4

COEFICIENTES DE FIABILIDAD α DE LAS SUBPRUEBAS DEL KM-R
CORRESPONDIENTES A LOS DIFERENTES CURSOS DE NUESTRA MUESTRA

	Coeficientes de Fiabilidad α											
	Preesc.	1°	2°	3°	CURSOS							1°BUP
				4°	5°	6°	7°	8°				
Subprueba 1	.70	.86	.88	.86	.87	.86	.81	.80	.86	.28	.61	
Subprueba 2	.62	.66	.80	.73	.82	.87	.87	.88	.89	.78	.85	
Subprueba 3	.76	.83	.87	.84	.80	.85	.82	.77	.83	.73	.90	
Subprueba 4	.78	.86	.83	.71	.92	.67	.66	.67	.64	.41	.61	
Subprueba 5	.61	.82	.85	.81	.91	.74	.70	.81	.74	.58	.68	
Subprueba 6	.56	.73	.91	.80	.73	.79	.76	.80	.77	.66	.73	
Subprueba 7	.59	.69	.96	.89	.81	.92	.79	.89	.85	.82	.93	
Subprueba 8	.52	.71	.83	.82	.85	.88	.84	.84	.83	.69	.89	
Subprueba 9	.79	.89	.83	.82	.93	.87	.77	.81	.82	.65	.77	
Subprueba 10	.32	.84	.89	.88	.96	.89	.87	.80	.87	.63	.80	
Subprueba 11	.49	.65	.72	.70	.81	.84	.86	.86	.81	.75	.66	
Subprueba 12	.66	.78	.83	.80	.89	.85	.84	.85	.84	.56	.70	
Subprueba 13	.58	.76	.75	.78	.84	.86	.83	.82	.82	.70	.78	
Área 1	.53	.69	.78	.66	.81	.80	.84	.77	.86	.76	.91	
Área 2	.68	.81	.88	.83	.78	.87	.83	.86	.86	.83	.72	
Área 3	.72	.86	.83	.85	.88	.90	.91	.88	.91	.86	.85	

(1992), se ha realizado un análisis de la correspondencia de los dominios del KM-R con el currículum del NCTM de 1989. Este trabajo ha mostrado que el KM-R es más adecuado que el resto de tests de matemáticas analizados para evaluar los contenidos, procesos y niveles de pensamiento determinados por los estándares de matemáticas del NCTM y, además, ha proporcionado un criterio de validez relativo a su adecuación de contenidos en su contexto de origen.

En segundo lugar, y con referencia al contexto educativo en el que la escala KM-R ha sido adaptada, se ha presentado un segundo estudio de Validación de Contenido, basado en el análisis de la correspondencia de los ítems al currículum, representado por unos modelos de Diseño Curricular en el ámbito de la aplicación de la LOGSE y por los Programas Renovados de la EGB. Los expertos seleccionados contrastaron los objetivos conductuales de cada uno de los ítems del KM-R con los objetivos de aprendizaje implícitos en cada uno de los dominios de los Diseños curriculares propuestos indicando, para cada una de las Subpruebas, el nivel escolar en el que se inicia la instrucción del objetivo y el nivel de logro de su dominio.

Los resultados de los análisis efectuados ponen de manifiesto la existencia de dicha adecuación curricular, confirmando la intencionalidad constructiva de la prueba y aportando una evidencia importante de validez de contenido y curricular. Además, la revisión lógica de la Validez Curricular ha servido de referente teórico para los análisis efectuados y para las interpretaciones del rendimiento, confirmándose, de nuevo, el KM-R como un instrumento para la evaluación y el diagnóstico.

En tercer lugar, se ha realizado un análisis de la Estructura Dimensional de la Escala con el fin de comprobar si las dimensiones obtenidas se corresponden con los núcleos de logro de los Diseños Curriculares correspondientes a nuestro análisis teórico. Mediante el Análisis Factorial de las Subpruebas para cada curso se han elaborado una serie de conclusiones referentes a la Validez de Constructo a partir de la adecuación de las definiciones de los dominios y representatividad de los ítems del KM-R. Estas conclusiones, respecto de la estructura factorial de los totales de las Subpruebas, destacan dos factores diferenciados marcados por la existencia o no de instrucción formal (ver Cuadro 1), y, en referencia a los dominios de

Cuadro 1

ESTRUCTURA FACTORIAL DOMINANTE Y SU DEFINICIÓN, RESULTANTE DEL ANÁLISIS FACTORIAL DE LOS TOTALES DE LAS SUBPRUEBAS DEL KM-R

ESTRUCTURA FACTORIAL DOMINANTE	
Factor 1.	Subpruebas con instrucción.
Factor 2.	Subpruebas de cálculo.
<i>Factor 3.</i>	<i>Subpruebas sin instrucción.</i>

las Subpruebas, las agrupaciones responden al nivel de logro de los estudiantes antes que a la congruencia de contenido como era esperable en una prueba de referencia curricular.

En cuarto lugar, se ha realizado un análisis de la Validez Criterial. Como criterio de contraste se ha construido una batería de pruebas que incluyen una serie de variables consideradas predictivas de rendimiento en matemáticas (ver Cuadro 2). Estas pruebas, junto al KM-R, se administraron a una muestra de sujetos representativa y correspondientes a tres niveles académicos. A partir de los resultados de la batería de pruebas, junto con los de escala KeyMath-R y mediante el análisis Cluster, se analizan los patrones diferenciados de ejecución para cada uno de los tres niveles de los sujetos de la muestra determinándose las características o perfiles de los grupos de

Cuadro 2

ÁMBITOS EN LOS QUE SE AGRUPAN LAS VARIABLES PREDICTIVAS DE RENDIMIENTO EN MATEMÁTICAS

- Aptitudes Mentales.
- Hábitos y Estilos Cognitivos.
- Factores de personalidad.
- Factores pedagógicos.
- Rendimiento anterior.
- Ambiente socio-económico.

estudiantes. Se ha comprobado que los perfiles de los grupos resultantes se deben al nivel de rendimiento a través de todas las Subpruebas del KM-R.

Asimismo, se ha realizado un análisis de varianza de una vía (Anova) con las variables criterio entre los grupos identificados, para determinar la significación de las diferencias entre ellos.

Posteriormente, a modo de un análisis clásico de Validez Convergente, se realiza un estudio correlacional de las variables del KM-R con las variables integradas en la batería de pruebas, para determinar la significación cualitativa de las relaciones observadas y mejorar el uso diagnóstico de la prueba.

Las Anovas se distribuyen entre los grupos según los patrones de rendimiento aunque solamente algunas de las variables relativas a la evolución académica e inteligencia guardan una relación significativa para todos los grupos por el rendimiento, pudiéndose señalar las siguientes cuestiones. El volumen inferior de correlaciones significativas corresponde a los grupos con un perfil de rendimiento en el KM-R más bajo. De entre las variables criterio correspondientes a la evolución académica que muestran correlaciones significativas con los grupos por rendimiento hay que señalar las correspondientes a las variables notas obtenidas en Matemáticas y en Lengua. También hay que destacar que, los grupos de más alto rendimiento dentro del grupo de sujetos de mayor nivel académico muestran correlaciones significativas a un nivel del 99% con algunas variables de la batería de Inteligencia utilizada. Estas últimas informaciones está siendo estudiada a un mayor nivel por cuanto permitirían considerar al KM-R como una prueba de rendimiento general con un alto valor predictivo.

Insistimos en el bajo volumen de correlaciones significativas encontrado entre las variables de la batería predictiva de rendimiento y los grupos por rendimiento en el KM-R, con excepción hecha de las correlaciones mencionadas anteriormente, y en el comportamiento prácticamente aleatorio de la distribución de las restantes correlaciones significativas. Por otra parte, se han elegido como variables criterio aquellas que la investigación ha ido considerando como predictivas de rendimiento en matemáticas.

Por todo ello, podemos afirmar que los resultados encontrados en este trabajo no corroboran las conclusiones de dichas investigaciones y que el perfil del estudiante de matemáticas encontrado es diferente al esperado.

Finalmente, y con el fin de cubrir todas las necesidades que comporta todo estudio de adaptación de una prueba de evaluación del rendimiento, se ha realizado un estudio diferencial del rendimiento como base a una posible baremación y, por último, ha realizado dicha baremación.

Como conclusión del trabajo de adaptación realizado, se puede afirmar que la Escala *KeyMath-Revised* de A. Connolly puede ser considerada una prueba general de matemáticas elementales referida al currículum y con una organización interna correspondiente al currículum. Ha mostrado unos niveles de bondad métrica muy aceptables, mejorando los estudios originales de Connolly, que avalan su uso; asimismo, teniendo en cuenta los resultados obtenidos en los diferentes estudios de validación, se revela como un instrumento diagnóstico de gran utilidad.

Además, la fundamentación teórica realizada a partir de algunas conclusiones del enfoque del Procesamiento de la Información y las revisiones lógicas de la Validez de Contenido realizadas, no existentes en la adaptación original de Connolly, otorgan al KM-R una mayor riqueza de interpretación diagnóstica y permiten diseñar programas de intervención basados en el propio currículum y en las conclusiones de la propia psicología cognitiva.

REFERENCIAS

- Bachor, D.G. (1989-90). *KeyMath-Revised (KMR)*. *Diagnostique*, vol. 15, nº 1-4, pp. 87-98.
- Connolly, A.J. (1985). *KeyMath Revised: teach and practice: Activities for the diagnosis and remediation of computation difficulties*. Circle Pines, Minnesota. American Guidance Service.
- Connolly, A.J. (1988). *KeyMath Revised: A Diagnostic Inventory of Essential Mathematics*. Circle Pines, Minnesota. American Guidance Service.
- Connolly, A.J. (1988a). *KeyMath Revised Assist. Automated System for Scoring and interpreting Standardized Tests*. Circle Pines, Minnesota. American Guidance Service.
- Mari, R. (1996). *Evaluación del rendimiento en matemáticas: estudio métrico diferencial de la escuela KeyMath-Revised*. Tesis Doctoral. Dep. MIDE, Universitat de València (inédito).
- National Council of Teachers of Mathematics (1980). *An agenda for action: Recommendations for school mathematics of the 1980's*. Reston, Virginia. Autor.
- National Council of Teachers of Mathematics (1989). *Curriculum and Evaluation Standards for School Mathematics*. (4ª Ed. 1991). Reston, Virginia. Autor.
- Pandey, T. (1994). *Authentic Mathematics Assessment*, Paper. California Department of Education. Obtenido por acceso mediante RED el 25 de mayo de 1994.
- Price, P.A. (1984). A comparative study of the California Achievement Test (Forms C and D) and the KeyMATH Diagnostic Arithmetic Test with secondary LH students. *Journal of Learning Disabilities*, vol. 17, nº 7, pp. 392-396.

- Romberg, T.A. (Ed.) (1992). *Mathematics assesment and evaluation. Imperatives for Mathematics Educators*. New York. State University of New York Press.
- Romberg, T.A., Wilson, L.D., M. Khaketla y Chavarria, S., (1992). Curriculum and test alignment. En Romberg, T.A. (Ed.), *Mathematics assesment and evaluation. Imperatives for Mathematics Educators*. New York. State University of New York Press, pp. 61-74.
- Tinney, F.A. (1975). A comparison of the KeyMath Diagnpstic Arithmetic Test used with learning disabled students. *Journal of Learning Disabilities*, vol. 8, n° 5, pp. 57-59.

UN EXAMEN DE LA VALIDEZ EXTERNA DEL STAT (STERNBERG TRIARCHIC ABILITIES TEST)

Ángela Rojo Martínez, M^a Dolores Prieto Sánchez* y Juan L. Castejón Costa**

* Universidad de Murcia, ** Universidad de Alicante

RESUMEN

El objetivo de nuestra investigación es comprobar la validez externa de tipo predictivo del STAT (Sternberg Triarchic Abilities Test). Para ello llevamos a cabo análisis de correlación y de regresión múltiple entre las puntuaciones obtenidas en el STAT, las conseguidas en el «g» de Cattell y un criterio externo: el rendimiento académico, en una submuestra de alumnos (n= 595) de 5º y 6º de E.G.B. elegida al azar. Los resultados de dichos análisis correlacionales y predictivos, evidencian que: 1) el STAT parece estar comprometido con el rendimiento académico en mayor medida que lo está el factor «g» de Cattell; 2) son la subpruebas analítica y práctica del STAT, las que tienen una mayor relación con el rendimiento; y 3) sin embargo, la subprueba creativa está menos relacionada con el rendimiento.

ABSTRACT

The aim of our investigation is to check the external and predictable validity of the STAT. To do we carried out the test of multiple correlation and drop between the marks obtained in the STAT, the ones in the «g» of Cattell and an external guideline: the academic performance in a subsample of pupils (n= 595) of the 5th and 6th years of Primary Education chosen at random. The results of these correlative and predictable tests demonstrate that: 1) The STAT seems to be linked to a greater extent with the academic performance than the factor «g» of Cattell is; 2) the analytical and practical subtests of the STAT have a greater connection with the performance; and 3) however, the creative subtest is connected to a lesser extent with the performance.

INTRODUCCIÓN

En la conceptualización de la inteligencia, la mayor parte de los investigadores sobre el tema, coinciden en afirmar el interés y la importancia que dicho fenómeno psicológico supone para la ciencia y para la sociedad. Esto ha llevado en los últimos cien años al estudio y desarrollo de diferentes conceptos alternativos sobre la inteligencia, que muestran ciertos núcleos básicos comunes, pero a su vez importantes puntos de discrepancia. Todo ello ha contribuido a impregnar de ciertas dosis de ambigüedad e inestabilidad la comprensión del mencionado constructo.

La evidencia de dichos cambios queda explicitada en la comparación que se establece entre los simposios de 1921 y 1986 sobre «La inteligencia y su medición», constatándose importantes variaciones en las prioridades de estudio temático. De tal forma que se ha pasado de una mayor atención prestada a cuestiones de conducta predictiva y pruebas psicométricas, en 1921, a una ampliación de intereses por el procesamiento de la información, por el papel del contexto y por las interacciones entre ambos, en 1986 (Sternberg y Detterman, 1992). De la comparación entre ambos simposios, se pueden deducir dos claras conclusiones: la primera, hace referencia al importante avance conseguido en la conceptualización del constructo de la inteligencia, llegándose en 1986 a definiciones más completas y elaboradas, donde la importancia social de la inteligencia y el interés de ésta por el «mundo real», ocupan un papel destacado. La segunda, explicita la necesidad de seguir ampliando el campo de la teoría y de la investigación de la inteligencia, para así obtener una mejor comprensión sobre la naturaleza de este escurridizo constructo.

Actualmente nos encontramos en un momento de cambios de enfoque y nuevos planteamientos, por la influencia de variables implícitas a la propia inteligencia y a su conceptualización, así como por la incidencia de aportaciones de carácter colateral especialmente de la Psicología Cognitiva y presiones desde los ámbitos aplicados, fundamentalmente los relacionados con la educación (Castelló, 1994).

Todo lo expuesto sobre el devenir de la historia de la inteligencia, nos lleva al estudio de las altas habilidades, ya que la inteligencia se ha considerado desde siempre como variable fundamental en la definición y evaluación de estos sujetos. De tal manera que se puede hablar de una estrecha relación entre el desarrollo de las teorías de inteligencia y la influencia de éstas en el concepto de superdotación; y a su vez la superdotación —como situación extrema de inteligencia— se convierte en fuente de interés teórica y de validación, para los modelos de inteligencia (Castelló, 1992).

Entre los teóricos cognitivos más destacados dentro del procesamiento de la información destacamos a Sternberg (1977), quien considera que la actuación inteligente se debe sobre todo a habilidades de tipo metacomponencial. Este autor pretende explicar el funcionamiento cognitivo y los mecanismos de autorregulación del superdotado desde un modelo dinámico, contextualizado en su Teoría Triárquica de la Inteligencia, base fundamental para entender la superdotación.

La Teoría Triárquica de la inteligencia, no es más que un caso especial de una teoría triárquica general, y como tal permite ofrecer una conceptualización de la inteligencia más amplia que la mayoría de las teorías convencionales (Sternberg, 1985). Y

puesto que la inteligencia, como dice Sternberg, no funciona en el vacío, sino en estrecha interacción con un mundo cuya complejidad es creciente, esto es lo que lleva a dicha teoría a explicar las relaciones entre: a) la inteligencia y el mundo interno del individuo, entendida ésta como los mecanismos mentales subyacentes en la considerada conducta inteligente, asimismo dicha inteligencia facilita la respuesta a la cuestión de cómo se genera el comportamiento (subteoría componencial o individual); b) la inteligencia y la experiencia, o aplicación de los mecanismos mentales para solucionar problemas que van desde los muy novedosos, como son los importantes procesos de «insight», hasta los muy familiares, lo que permite poner en relación la inteligencia con el mundo interno y externo, y responder a la pregunta de cuándo el comportamiento es inteligente (subteoría experiencial); y c) la inteligencia y el mundo externo del individuo o uso de esos mecanismos mentales para lograr la adaptación al medio, y a su vez plantear cuestiones de qué comportamientos son inteligentes, para quién y dónde (subteoría contextual).

Este autor denomina «triárquica» a su teoría porque utiliza las tres subcategorías o subteorías mencionadas para explicar la superdotación. Estableciendo desde la globalidad de la teoría, una estructura jerárquica caracterizada por la división en las subteorías indicadas, que a su vez se organizan en grupos sucesivamente menores, sin que por ello pierdan la estrecha interconexión entre las mismas. De forma tal que las distintas subteorías, en razón a sus funciones, quedan concretadas en: 1) la componencial o individual, definida por los metacomponentes, los componentes de rendimiento y los de conocimiento-adquisición; 2) la experiencial o creativa, especificada por la capacidad de insight en la solución de tareas novedosas y la automatización de la información; y 3) la contextual o práctica, entendida a través de tres mecanismos referidos a la adaptación, la configuración y la selección de los ambientes del mundo real, que resultan relevantes para la vida del sujeto.

En suma, la teoría triárquica de Sternberg pretende explicar los mecanismos cognitivos, mediante los cuales el individuo procesa y automatiza la información, para así conseguir la adaptación al medio social donde se desarrolla. Desde esta teoría la superdotación se plantea como algo complejo, de naturaleza plural y múltiple en sus manifestaciones; convirtiéndose dicha teoría en uno de los modelos más actuales en el estudio de la superdotación.

Así, Sternberg (1991) con el diseño del STAT (Sternberg Triarchic Abilities Test, 1991) pretende que éste sirva como instrumento de evaluación de los procesos y funciones de las subteorías que componen la teoría triárquica (subteoría componencial, experiencial y contextual); considerando fundamental en dicho proceso tanto el contexto, como la experiencia del individuo, y de ahí el estimar la influencia cultural presente en todo tipo de test. Por tanto, el STAT ha supuesto una alternativa innovadora como modelo de evaluación de la inteligencia, ayudando especialmente a la determinación de las diferencias individuales entre los sujetos con altas habilidades y los de capacidades medias.

La prueba en su versión original consta de un total de 90 preguntas, organizadas en 9 escalas y agrupadas en tres categorías, que a su vez pretenden evaluar los diferentes tipos de inteligencia, en las tres modalidades (verbal, numérica y figurativa) emplea-

das en el procesamiento de la información. La primera categoría agrupa las pruebas 1, 2 y 3 referidas a los «metacomponentes individuales» de la inteligencia, en sus respectivas modalidades. La segunda categoría se concreta en la «inteligencia práctica» e incluye las pruebas 4, 5 y 6, igualmente en las tres modalidades indicadas. Y la tercera categoría está referida al aspecto «experencial» y de «insight» de la inteligencia, en sus correspondientes modalidades —pruebas 7, 8 y 9—; además en esta categoría se sitúa también una prueba de «automatización» —prueba 10—.

Siendo así, que para la evaluación de la «inteligencia individual» se utilizan tareas en las que se incluyen problemas relacionados con los componentes o habilidades intelectuales académicas. Mientras que la «inteligencia experencial» se valora mediante tareas cuya solución exige de la aplicación de esa forma nueva de pensamiento y desde ésta se considera especialmente todo el proceso de recomposición seguido por el individuo, más que la respuesta novedosa en sí; dentro de esta inteligencia se sitúa también la escala número 10, anteriormente citada, y que como prueba de «automatización» resulta la excepción del STAT, al contemplar los ítemes que la forman la rapidez del sujeto para realizar una determinada tarea. Y la tercera de las categorías, conocida como «inteligencia práctica o contextual» se mide con problemas que requieren la aplicación de procesos de razonamiento lógico, relacionados estrechamente con la vida práctica (Galindo, Prieto y Rojo, 1995).

En definitiva, el objetivo fundamental de nuestro trabajo de investigación se centra en el estudio de la validez externa —de carácter predictivo— del STAT (Sternberg Triarchic Abilities Test o Test de Habilidades de la Inteligencia Triárquica), tomando como criterio externo el rendimiento académico. Aunque en general, la validación del STAT plantea diversas cuestiones de carácter teórico, al resultar ser una prueba surgida de un planteamiento teórico amplio sobre la inteligencia, y a su vez, tratarse de una prueba más de inteligencia con matices diferenciadores, de ahí que era de esperar que mostrara no sólo una validez convergente, sino también discriminante con otras pruebas utilizadas tradicionalmente.

Dado que los procedimientos seguidos para la validación de la prueba han incluido los distintos tipos de validez (interna; externa; estructural y teórico-experimental), hemos preferido en el presente trabajo profundizar en la validación externa, puesto que tanto la interna, como la estructural y la teórico-experimental han sido y están siendo motivo de otras investigaciones (Bermejo, 1995; Bermejo y Castejón, 1995).

METODOLOGÍA

Sujetos

La muestra de sujetos está compuesta por 1.255 alumnos/as de 5^o y 6^o cursos de E.G.B. pertenecientes a 22 Centros de Enseñanza General Básica de la Región de Murcia, entre los cuales se extrae una submuestra de 595 estudiantes, elegidos al azar, pero con suficiente representatividad, tanto en lo que respecta a la muestra total, como a la población de referencia.

Instrumentos

Los instrumentos utilizados incluyen las pruebas de inteligencia de factor «g» de Cattell y la prueba STAT de Sternberg (*Sternberg Triarchic Abilities Test*).

El test de inteligencia general, *factor «g» de Cattell* (Cattell y Cattell, 1973) es una prueba que se considera en buena medida «libre de influencias culturales»; siendo uno de los instrumentos más utilizados en la evaluación de la inteligencia en el nivel de edad en el que trabajamos.

Como es conocido, la mencionada prueba consta de cuatro subpruebas: Series, Clasificaciones, Matrices y Condiciones; que ofrecen una puntuación total y una puntuación CI de inteligencia general.

Aunque el test se encuentra adaptado en nuestro país, el tamaño de la muestra utilizada en este estudio nos ha permitido llevar a cabo la normalización de la prueba por separado, para cada uno de los cursos 5º y 6º; encontrando unos resultados similares a los que se ofrecen en el manual de la adaptación española, con unos baremos equivalentes en ambos cursos.

La *prueba STAT (Sternberg Triarchic Abilities Test)* de R.J. Sternberg (1991) es un instrumento de evaluación de la capacidad intelectual que se encuentra en fase de experimentación; y que como ya se explicó con anterioridad, surge a partir de la elaboración teórica del autor sobre las características de la inteligencia de los sujetos con altas habilidades, aunque su ámbito de aplicación incluye la población normal. El test se compone de 9 escalas, que contienen cada una 10 preguntas o ítems, más dos ejemplos resueltos; esta prueba puede administrarse en grupo y existen diferentes versiones experimentales de la misma para los distintos niveles de edad, desde los 4 años hasta la edad adulta o «college». La versión utilizada en este trabajo va destinada al Nivel E, que comprende a los niños situados entre las edades de 10 a 11 años de nuestro sistema educativo.

Además de evaluar con estos instrumentos las variables referidas al Cociente Intelectual General, medido por el factor «g» de Cattell y el STAT de Sternberg, también se incluye otra variable relacionada con las calificaciones en las materias de Lengua Castellana, Matemáticas, Ciencias Naturales y Ciencias Sociales.

Procedimiento

Inicialmente se procede a la aplicación y adaptación de la prueba de inteligencia STAT, y la prueba de factor «g» de Cattell. Una vez seleccionados los sujetos que constituyen la muestra inicial, se lleva a cabo la aplicación del test de factor «g» de Cattell a los alumnos de 5º y 6º cursos de EGB. La aplicación de esta prueba se realiza durante el curso, en las aulas de clase, aprovechando el horario escolar, y según las instrucciones contenidas en el manual del test.

En lo que respecta al STAT, al tratarse de una «prueba de potencia», en las instrucciones para la realización de ésta, se indica a los sujetos que no hay tiempo límite para su cumplimentación, si bien han de llevarla a cabo sin perder demasiado tiempo. Asimismo, se les informa, que la prueba no tiene transcendencia para sus calificaciones,

porque se trata de ver cómo se utilizan las estrategias de pensamiento. En las instrucciones que se dan a los alumnos, la persona encargada de aplicar las pruebas explica detalladamente la realización de los dos ejemplos resueltos que están al inicio de cada subprueba, asegurándose que han sido entendidos por toda la clase. La aplicación del STAT tiene lugar durante dos sesiones matutinas, de aproximadamente una hora cada una; en cada sesión se realiza la mitad del test, o lo que es lo mismo 5 subpruebas.

Tras la aplicación de las pruebas mencionadas, se recogen también las calificaciones medias de la submuestra de sujetos elegidos al azar, en las materias con anterioridad indicadas.

RESULTADOS

En primer lugar, conviene indicar que puesto que una de las pruebas de evaluación de la habilidad intelectual se encuentra aún en fase de experimentación, el Sternberg Triarchic Abilities Test, resulta necesario proceder a adaptar esta prueba en el curso de esta investigación.

La adaptación de cualquier prueba de evaluación y medida en el ámbito de las capacidades, requiere la comprobación de la bondad de unas características psicométricas mínimas, que se cifran en la traducción de la prueba, una aplicación «piloto», el análisis de ítemes, la baremación, fiabilidad y validez. Sin embargo, preferimos en el caso del STAT de Sternberg, únicamente nombrar los citados pasos al existir trabajos que abordan de forma específica la adaptación de la prueba (Galindo, Martínez y Prieto, 1996).

De esta forma centramos el análisis de resultados en la finalidad principal de este estudio, como es el examen de la validez externa de tipo predictivo del Sternberg Triarchic Abilities Test, utilizando como criterio el rendimiento académico.

Análisis correlacional

Para la comprobación del grado de validez externa del STAT, hemos llevado a cabo análisis de correlación y de regresión múltiple entre las puntuaciones obtenidas en la prueba y el rendimiento académico, tomado como criterio externo; en una submuestra de sujetos elegida al azar ($n= 595$), pero suficientemente representativa de la muestra total y de la población de referencia.

Los resultados del análisis de correlación entre las puntuaciones de Cociente Intelectual del factor «g», del STAT, y a su vez, las distintas subpruebas del Sternberg Triarchic Abilities Test, con el rendimiento académico, definido operativamente por las notas escolares, se ofrecen en la tabla 1.

Como podemos observar en la tabla de correlaciones, las puntuaciones de Cociente Intelectual en el STAT mantienen una correlación mayor con el rendimiento académico ($r= .50$), que las puntuaciones de Cociente Intelectual en la prueba de factor «g» ($r= .40$). El STAT parece estar más comprometido en la evaluación de aspectos relacionados con el rendimiento, que la prueba de factor «g». Son las subpruebas analítica y práctica del STAT, las que obtienen una mayor relación con el rendimiento ($r=.46$ en

Tabla 1

	CIG	CIST	ANALIT	PRACTI	CREATI	PRUEBA1	PRUEBA2	PRUEBA3	PRUEBA4	PRUEBA5	PRUEBA6	PRUEBA7	PRUEBA8	PRUEBA9	NOTAMED
CIG	1.00														
CIST	.58**	1.00													
ANALIT	.50**	.80**	1.00												
PRACTI	.52**	.85**	.64**	1.00											
CREATI	.47**	.86**	.56**	.60**	1.00										
PRUEBA1	.30**	.44**	.57**	.37**	.31**	1.00									
PRUEBA2	.39**	.68**	.82**	.54**	.48**	.26**	1.00								
PRUEBA3	.42**	.64**	.80**	.50**	.43**	.24**	.46**	1.00							
PRUEBA4	.33**	.58**	.42**	.74**	.35**	.26**	.32**	.34**	1.00						
PRUEBA5	.43**	.67**	.56**	.79**	.48**	.36**	.49**	.40**	.41**	1.00					
PRUEBA6	.42**	.69**	.48**	.76**	.53**	.22**	.42**	.40**	.31**	.40**	1.00				
PRUEBA7	.22**	.49**	.26**	.31**	.63**	.19**	.20**	.20**	.23**	.21**	.27**	1.00			
PRUEBA8	.43**	.75**	.529**	.56**	.82**	.29**	.49**	.35**	.31**	.47**	.50**	.23**	1.00		
PRUEBA9	.38**	.65**	.43**	.42**	.73**	.18**	.35**	.39**	.22**	.36**	.38**	.18**	.49**	1.00	
NOTAMED	.40**	.50**	.46**	.46**	.34**	.34**	.38**	.32**	.34**	.39**	.32**	.13**	.36**	.23**	1.00

Nº de casos: 595

1-cola Signif: * — .01

** — .001.

CIG= Cociente intelectual en factor «G»; CIST= Cociente intelectual en el STAT; ANALIT= Puntuaciones en la parte analítica del STAT; PRACTI= Puntuaciones en la parte práctica del STAT; CREATI= Puntuaciones en la parte creativa del STAT; NOTAMED= Nota media del rendimiento académico.

ambos casos), mientras que la subprueba creativa parece medir aspectos distintos al rendimiento académico ($r = .34$). La subprueba de STAT considerada como creativa, es asimismo la que mantiene una relación más baja con el Cociente Intelectual obtenido en el factor «g».

En general los estudios de validación preliminar del STAT han mostrado que esta prueba está relacionada con otros test como el factor «g» de Cattell. Aunque el porcentaje de varianza que comparten es lo suficientemente moderado como para no considerar al STAT una prueba más de inteligencia general de tipo abstracto, como lo es el test de factor «g». Así, Sternberg y Clinkenbeard (1995) encuentran las siguientes correlaciones con la prueba de Cattell, .50 para la subprueba Analítica, .55 para la Creativa y .36 para la Práctica.

Por otra parte, la correlación entre las subpruebas del STAT también es moderada, destacando el valor de .64 entre las subpruebas Analítica y Práctica. Los primeros resultados ofrecidos por Sternberg y colaboradores (Sternberg, Ferrari, Clinkenbeard y Grigorenko, 1996) muestran unos índices de correlación que se sitúan en .23 entre las subpruebas Analítica y Creativa, y .14 entre la Analítica y la Práctica. Aunque estos valores son más bajos, hay que señalar que los resultados que ofrecen los autores citados se obtienen con un nivel distinto de la misma prueba. El nivel H, destinado a los alumnos de los últimos cursos de educación secundaria y bachillerato.

La relación que mantienen las subpruebas del STAT con el rendimiento, expresado en la calificación media en las cuatro asignaturas, se sitúa en un rango medio, pues oscila entre .46 con la subprueba Analítica y .34 con la Creativa. Los valores obtenidos por Sternberg y colaboradores (1996) con el nivel H se sitúan alrededor de .60.

Análisis de regresión múltiple: método paso a paso (stepwise)

Los resultados del análisis de regresión múltiple, que presentamos en la tabla 2, toman como variables predictoras las puntuaciones en la prueba de factor «g» de Cattell, y cada una de las tres subpruebas del STAT —la analítica, la práctica y la creativa—, y como variable criterio las calificaciones medias en Lengua Castellana, Matemáticas, Ciencias Naturales y Ciencias Sociales. Se emplea el método paso a paso (stepwise), para la selección de las variables que entran en la ecuación.

Las variables que logran predecir en mayor medida el rendimiento son la capacidad analítica y la práctica, evaluadas por el STAT y la inteligencia medida por el factor «g» de Cattell, por este orden. El valor de la correlación múltiple es moderado ($R = .53$), lo que logra explicar el 28 por ciento de la varianza observada en el rendimiento.

De nuevo se observa que son los aspectos analíticos y prácticos de la inteligencia, medidos por el STAT, los que tienen un mayor compromiso con el rendimiento académico. El factor «g» no obstante sigue mostrando una contribución independiente, lo que nos indica nuevamente, que a pesar de la existencia de componentes comunes en la realización de una y otra prueba, parece haber algunos aspectos específicos recogidos en cada una de ellas de forma independiente.

Además los resultados del análisis de regresión múltiple paso a paso obtenidos por Sternberg y colaboradores (Sternberg et al., 1996) para el nivel H del STAT, cuando se

Tabla 2

RESUMEN DEL ANÁLISIS DE REGRESIÓN MÚLTIPLE, CON LAS SUBPRUEBAS DEL STAT Y EL COCIENTE INTELLECTUAL DEL FACTOR «G», COMO PREDICTORES Y EL RENDIMIENTO ACADÉMICO COMO CRITERIO

R Múltiple			.53080		
R Cuadrado			.28175		
R Cuadrado Ajustada			.27811		
Error Estándar			1.02407		
Análisis de Varianza					
	GL	Suma de Cuadrados	Media Cuadrados		
Regression	3	243.12964	81.04321		
Residual	591	619.78779	1.04871		
F =	77.27893	Signif F =	.0000		
VARIABLES EN LA ECUACIÓN					
Variable	B	ES B	Beta	T	Sig T
ANALIT	.074469	.014976	.234967	4.972	.0000
PRACTI	.060213	.012859	.223753	4.683	.0000
CIG	.014391	.003609	.169133	3.988	.0001
(Constant)	-1.119656	.346851	-3.228	.0013	

toman como predictores las subpruebas del STAT, y como variable criterio el rendimiento académico muestra un coeficiente de correlación múltiple R de .45; semejante al encontrado por nosotros. Los coeficientes de regresión estandarizados β son de .20 para la subprueba Analítica, .18 para la Creativa y .16 para la Práctica. Nuestros resultados indican, sin embargo, un mayor peso para la Práctica que para la Creativa, que no hace una contribución significativa a la explicación del rendimiento en ninguno de los casos.

En la tabla 3 aparecen los resultados de los análisis de regresión paso a paso sobre cada una de las calificaciones por separado, tomando como variables predictoras cada una de las subpruebas del STAT y la puntuación alcanzada en la prueba de factor «g».

Como podemos observar las variables que hacen una contribución significativa a la explicación de cada una de las calificaciones son, en todos los casos, las subpruebas

Tabla 3

RESULTADOS DEL ANÁLISIS DE REGRESIÓN, SIGUIENDO EL MÉTODO PASO A PASO (STEPWISE) DE LAS SUBPRUEBAS DEL STAT SOBRE LAS CALIFICACIONES EN LAS DISTINTAS ASIGNATURAS

LENGUA					
<i>Variable</i>	<i>B</i>	<i>SEB</i>	β	<i>T</i>	<i>SigT</i>
ANALIT	.0765	.0174	.2154	4.39	.0000
PRACTI	.0573	.0149	.1901	3.83	.0001
CIG	.0152	.0041	.1596	3.63	.0003
R múltiple= .4772; Sig= .0000					
MATEMÁTICAS					
<i>Variable</i>	<i>B</i>	<i>SEB</i>	β	<i>T</i>	<i>SigT</i>
ANALIT	.0881	.0162	.2557	5.41	.0000
PRACTI	.0605	.0139	.2069	4.33	.0000
CIG	.0151	.0039	.1635	3.85	.0001
R múltiple= .5303; Sig= .0000					
CC. NATURALES					
<i>Variable</i>	<i>B</i>	<i>SEB</i>	β	<i>T</i>	<i>SigT</i>
PRACTI	.0591	.0146	.2039	4.04	.0001
ANALIT	.0625	.0170	.1829	3.66	.0003
CIG	.0129	.0041	.1410	3.15	.0017
R múltiple= .4466; Sig= .0000					
CC. SOCIALES					
<i>Variable</i>	<i>B</i>	<i>SEB</i>	β	<i>T</i>	<i>SigT</i>
PRACTI	.0636	.0139	.2226	4.55	.0000
ANALIT	.0705	.0162	.2095	4.33	.0000
CIG	.0142	.0039	.1575	3.63	.0003
R múltiple= .4987; Sig= .0000					

Analítica y Práctica del STAT y el CI obtenido en el test de factor «g». Por el contrario, la subprueba Creativa del STAT no explica parte de la varianza en ninguna de las calificaciones en lo referido a cualquiera de las cuatro asignaturas.

Análisis de regresión múltiple: método de análisis jerárquico

La técnica de análisis de regresión paso a paso (stepwise) es útil para establecer la contribución única de cada una de las variables predictoras, sobre todo cuando las variables predictoras se encuentran correlacionadas entre sí. Su empleo resulta más adecuado en modelos predictivos.

Por contra, el método de partición incremental de la varianza o de análisis jerárquico de regresión supone partir de un modelo teórico acerca de la influencia causal de las variables. El procedimiento, como es conocido, consiste en ir añadiendo sucesivamente variables a la ecuación de predicción a la vez que se examina la significación del incremento en el coeficiente de regresión múltiple que aporta cada variable. En este procedimiento, las variables que entran primero en la ecuación explican toda la varianza que se debe a sí mismas y la contribución común que comparten con otras variables que entran posteriormente en la ecuación de regresión. De forma que cada variable adicional incluye la varianza que ésta tiene en común con las variables que se incluyen después. Solamente la última variable introducida en la ecuación contribuye de forma única a la varianza explicada. Por consiguiente, la decisión de introducir primero unas u otras variables debe hacerse sobre bases teóricas.

La tabla 4 muestra los resultados del análisis de regresión jerárquica cuando se introducen de forma sucesiva en la ecuación las puntuaciones correspondientes al factor «g» (CIG) y al STAT (CIST). En ambos casos la contribución de la segunda variable lleva a un incremento considerable en la explicación del rendimiento, poniendo así de manifiesto que se trata de pruebas distintas, que sólo comparten una proporción baja de varianza.

Tabla 4

RESULTADOS DEL ANÁLISIS JERÁRQUICO DE REGRESIÓN REALIZADO SOBRE EL RENDIMIENTO CON LAS DOS PRUEBAS DE NIVEL INTELECTUAL EL CI EN FACTOR «G» Y EL CI DEL STAT

Paso	Variable	R	R ² ajust.	R ² cambio	Fcambio	Sig.F
1	CIG	.40	.16	.16	116.94	.0000
2	CIST	.52	.27	.10	89.09	.0000
1	CIST	.50	.25	.25	203.30	.0000
2	CIG	.52	.27	.01	15.23	.0000

N= 594. CIG= Cociente intelectual en el factor «g»; CIST= Cociente intelectual en el STAT.

En la tabla 5 aparecen los resultados del análisis cuando se tienen en cuenta como variables explicativas las diferentes subpruebas del STAT. La explicación del rendimiento se ve incrementada por la inclusión de las subpruebas del STAT, más allá de la ofrecida por el CI obtenido en el factor «g» que se introduce en primer lugar.

Tabla 5

RESULTADOS DEL ANÁLISIS JERÁRQUICO DE REGRESIÓN REALIZADO SOBRE EL RENDIMIENTO CUANDO SE INCLUYEN DE FORMA SUCESIVA LAS VARIABLES CIG Y LOS FACTORES ANALÍTICO, PRÁCTICO Y CREATIVO DEL STAT

Paso	Variable	R	R ² ajust.	R ² cambio	Fcambio	Sig.F
1	CIG	.40	.16	.16	116.94	.0000
2	ANALIT	.50	.25	.09	71.83	.0000
3	PRACTI	.53	.27	.02	21.92	.0000
4	CREATI	.53	.27	.00	.05	.8077

N= 594. CIG= Cociente intelectual en el test de factor «g»; ANALIT= Subprueba Analítica del STAT; PRACTI= Subprueba Práctica del STAT; CREATI= Subprueba Creativa del STAT.

Son las subpruebas Analítica y Práctica del STAT las que hacen una contribución significativa a la explicación del rendimiento académico. La subprueba Creativa no tiene un efecto significativo.

Para comprobar si se produce la contribución de la subprueba Creativa cuando se incluye primero en la ecuación, se introducen por este orden las subpruebas Creativa, Práctica y Analítica del STAT, y a continuación la puntuación en el factor «g». Aunque la subprueba Creativa tiene un efecto significativo ($\beta=.34$; $T=8.83$; $Sig=.0000$) cuando se introduce en primer lugar, este efecto disminuye al incluir la subprueba Práctica ($\beta=.09$; $T=2.11$; $Sig=.0345$) y deja de ser claramente significativo cuando se introduce la subprueba Analítica ($\beta=.021$; $T=.45$; $Sig=.6475$). Los resultados finales de la ecuación de regresión cuando se incluyen todas las variables mencionadas son los que se ofrecen en la tabla 6.

Estos resultados, junto a los de la tabla anterior indican que la subprueba Creativa del STAT tiene elementos comunes con las subpruebas Práctica y Analítica, sobre todo con esta última, cuando se trata de explicar el rendimiento académico.

DISCUSIÓN

Primeramente conviene indicar que los resultados, tomados en su conjunto, ponen claramente de manifiesto que el Sternberg Triarchic Abilities Test (STAT) no es una medida más de la capacidad general o del factor «g».

Tabla 6

RESULTADOS DEL ANÁLISIS DE REGRESIÓN JERÁRQUICA, TOMADO COMO CRITERIO EL RENDIMIENTO MEDIO, DE LAS VARIABLES DE INTELIGENCIA CREATIVA, PRÁCTICA Y ANALÍTICA DEL STAT, ASÍ COMO EL CIG

Múltiple R	.53087				
R Cuadrado	.28183	Cambio R Cuadrado	.01914		
R Cuadrado Ajustado	.27696	Cambio en F	15.72133		
Error Standard	1.02488	Signif Cambio F	.0001		
Análisis de Varianza					
	GL	Suma de Cuadrados	Media Cuadrados		
Regression	4	243.19195	60.79799		
Residual	590	619.72549	1.05038		
F =	57.88178	Signif F =	.0000		
VARIABLES EN LA ECUACIÓN					
Variable	B	ES B	Beta	T	Sig T
CREATI	-.002502	.010275	-.011324	-.244	.8077
PRACTI	.061319	.013646	.227862	4.494	.0000
ANALIT	.075351	.015420	.237750	4.887	.0000
CIG	.014548	.003669	.170986	3.965	.0001
(Const.)	-1.132034	.350828	-3.227	.0013	

N= 594. CREATI= Inteligencia creativa del STAT. PRACTI= Inteligencia práctica del STAT. ANALIT= Inteligencia analítica del STAT. CIG= Cociente intelectual en el factor «g».

En consonancia con el objetivo fundamental propuesto en este trabajo, dirigido al estudio de la validez externa del Sternberg Triarchic Abilities Test, los resultados de los análisis correlacionales y predictivos, tomando como criterio el rendimiento académico, evidencian que el STAT parece encontrarse en mayor medida comprometido con el rendimiento académico, que lo que muestra estarlo el factor «g» de Cattell.

Por otra parte, la baja relación entre la subprueba creativa del STAT con el Cociente Intelectual obtenido en el factor «g», puede indicarnos, que es esta parte de la inteligencia medida por el STAT, la que tiene unos componentes más específicos, y en

buena medida, distintos tanto al rendimiento académico, como a la inteligencia general de tipo analítico y abstracto, como es la medida por el factor «g».

En lo que respecta a las relaciones que se producen entre las nueve subpruebas en que se divide el STAT, se observan correlaciones entre moderadas y altas entre todas las pruebas, sin que por ello aparezca un patrón claro de correlaciones positivas y altas entre las subpruebas que miden teóricamente un mismo aspecto de la inteligencia (analítico, práctico o creativo) y correlaciones negativas o moderadas entre subpruebas que miden aspectos diferentes, conforme a lo esperado teóricamente.

Si atendemos a la diferenciación entre una inteligencia fluida y una inteligencia cristalizada (Horn, 1994), la nueva prueba de inteligencia parece estar más comprometida con esta última. El componente específico del STAT parece encontrarse en la habilidad del sujeto para manejar la información en contexto (Sternberg, 1995), más allá de las capacidades de memoria y razonamiento abstracto exigidas por las pruebas de factor «g».

Cuando se realiza una partición de la varianza que separa los componentes específicos implicados en la realización de la prueba de factor «g», de los componentes cognitivos que exige la realización del STAT, se observa claramente que esta prueba aporta una contribución significativa a la explicación del rendimiento académico.

La subprueba Creativa del STAT, sin embargo, es la menos implicada en la explicación del rendimiento. Esto puede deberse a dos causas alternativas; una, que la subprueba Creativa del STAT tenga verdaderamente elementos comunes con las subpruebas Práctica y Analítica, sobre todo con esta última, cuando se trata de explicar el rendimiento académico; otra, que el rendimiento al final de la educación primaria e inicio de la secundaria esté determinado en mucha mayor medida por los componentes ejecutivos, analíticos y prácticos de la inteligencia, más que por los factores creativos, cuya influencia se dejaría sentir en cursos superiores cuando ya se dominan los componentes anteriores, o incluso en la actividad académica o profesional mucho más avanzada, tal como en cursos de especialización o doctorado, en la propia actividad profesional, etc., como ha sugerido el propio autor (Sternberg, 1985; 1995).

En todo caso, el STAT aparece como una nueva prueba de evaluación de la inteligencia con componentes específicos que no se solapa con otras pruebas de inteligencia general; así como un nuevo instrumento útil para la predicción y/o explicación de criterios externos tales como el del rendimiento académico.

Finalmente, conviene indicar que aunque los procedimientos de validación de la prueba, muestran resultados bastantes coherentes, consideramos necesario seguir profundizando en criterios externos de validación; pues la definición de criterios externos de competencia es una de las cuestiones que permanece abierta en la identificación de sujetos con altas habilidades. De ahí la conveniencia para una mejor comprensión de los procesos que se ponen en marcha para responder de forma adecuada a las tareas que contiene el STAT, continuar sometiendo a esta prueba a procedimientos de validez externa en los que se utilicen diferentes criterios de validación, conjuntamente considerados.

REFERENCIAS BIBLIOGRÁFICAS

- Bermejo, M.R. (1995). *El insight en la solución de problemas: Cómo funciona los superdotados*. Tesis Doctoral. Universidad de Murcia.
- Bermejo, M.R. y Castejón, J.L. (1995). Mecanismos de selección de la información en sujetos superdotados. *FAISCA*, nº 3, 45-58.
- Castelló, A. (1992). Concepto de superdotación y modelos de inteligencia. En Y. Benito (Coord.). *Desarrollo y educación de los niños superdotados* (pp. 19-35). Salamanca: Amarú.
- Castelló, A. (1994). Inteligencia Humana. En Y. Benito (Coord.). *Intervención e investigación psicoeducativas en alumnos superdotados* (pp. 17-25). Salamanca: Amarú.
- Cattell, R.B. y Cattell, A.K.S. (1974). *Tests de factor «g». Escalas 2 y 3*. Madrid: TEA.
- Galindo, A., Prieto, M.D. y Rojo, A. (1995). La inteligencia triárquica en el estudio de los superdotados. *Comunicación presentada al II Congreso Internacional de Psicología y Educación*. Celebrado en Madrid los días 16 a 18 de Noviembre de 1995.
- Galindo, A., Martínez, F. y Prieto, M.D. (1996). Adaptación del STAT (Sternberg Triarchic Abilities Test). *FAISCA*, nº 4, 38-53.
- Horn, J.L. (1994). Theory of fluid and crystallized intelligence. En R.J. Sternberg (Ed.), *The encyclopedia of human intelligence*. (Vol. 1, pp. 443-451). Nueva York: Macmillan.
- Rojo, A. (1996). *La identificación de alumnos con altas habilidades: enfoques y dimensiones actuales*. Tesis Doctoral. Universidad de Murcia.
- Sternberg, R.J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Hillsdale, N.J.: Erlbaum.
- Sternberg, R.J. (1985). *Beyond IQ. A triarchic theory of human intelligence*. Cambridge, MA: Cambridge University Press (Traduc. cast., 1990. *Más allá del C.I.*, Bilbao: DDB).
- Sternberg, R.J. (1991). Theory-based testing of intellectual abilities rationale for the triarchic abilities test. En H.A. Rowe (Ed.), *Intelligence: Reconceptualization and measurement*. Hillsdale, NJ: LEA.
- Sternberg, R.J. (1995). *In search of human mind*. Orlando: Harcourt Brace & Javanovich.
- Sternberg, R.J. y Detterman, D.K. (1986) (Comp.). *What is Intelligence?* New Jersey: Ablex (Traduc. cast., 1992. *¿Qué es la inteligencia? Enfoque actual de su naturaleza y definición*. Madrid: Pirámide).
- Sternberg, R.J.; Ferrari, M. Clinkenbeard, P. y Grigorenko, E. (1996). Identification, Instruction, and Assessment of Gifted Children: A Construct Validation of a Triarchic Model (En prensa).
- Sternberg, R.J. and Clinkenbeard, P. (1995). A triarchic view of identifying, teaching, and assessing gifted children. *Roeper Review*, 17, (4), 255-260.

DETERMINANTES DEL PROGRESO EN LOS ESTUDIOS DE DOCTORADO. DIFERENCIAS ENTRE DISCIPLINAS CIENTÍFICO-TECNOLÓGICAS VERSUS CIENCIAS SOCIALES Y HUMANIDADES*

Esperança Villar**
Universitat de Girona

RESUMEN

En numerosos países, incluido España, existen problemas en relación con las tasas de éxito en los estudios de doctorado. El trabajo que aquí se presenta fue diseñado con el fin de identificar los factores asociados con el progreso durante el tercer ciclo y la elaboración de las tesis doctorales en dos grupos disciplinares: científico-tecnológico versus humanidades y ciencias sociales.

El artículo presenta la información obtenida a partir de un cuestionario aplicado a 81 estudiantes de doctorado de la Universidad de Girona, en relación con tres áreas de interés: variables económicas, personales y académicas que condicionan el éxito en el tercer ciclo.

Los resultados muestran diferencias substanciales entre disciplinas. Los factores responsables de estas diferencias son, principalmente: la vinculación profesional a la universidad, la planificación del proyecto de investigación y los objetivos personales.

ABSTRACT

Problems in doctoral education, especially with respect to completion rates, have emerged in many countries, including Spain. This study was designed to identify the factors associated with the progress to the doctoral degree and the elaboration of doctoral dissertations in two disciplinary groups: hard sciences and technology versus humanities and social sciences.

* Esta investigación ha sido subvencionada por la Caixa de Girona.

** Departamento de Psicología. Plaça Sant Domènec, 9. 17071 Girona (evillar@zeus.udg.es)

This paper presents information from a questionnaire completed by 81 doctoral students in the University of Girona, attending three main areas of inquiry: economic, personal and academic variables conditioning outcomes.

Results show substantial differences between disciplines. The most important factors accounting for the differences are: professional involvement in the university, planning of the doctoral research project and personal goals.

Some considerations for enhancing doctoral education are given.

INTRODUCCIÓN

Ha sido, en general, escaso el interés suscitado en España por el tercer ciclo universitario como objeto de investigación psicopedagógica, teniendo en cuenta su trascendencia en relación con los dos ejes fundamentales de la actividad académica: la investigación —a través de la capacitación de los nuevos científicos y la consolidación de equipos de trabajo—, y la docencia —mediante la formación del futuro profesorado universitario—. Esta indiferencia ha corrido paralela a la falta de medios materiales y al «escaso control e interés de que han sido objeto estos estudios» por parte de la Administración, tal como se reconoce en el Real Decreto 185/85 que los regula.

Aunque esta situación no es exclusiva de la Universidad española, se dispone ya de cierto número de trabajos sobre el tercer ciclo publicados durante la última década. Uno de los temas que con mayor insistencia han abordado los especialistas en este ámbito es, sin duda alguna, el escollo que representa la realización de la tesis doctoral. Dos datos suelen citarse como ejemplo de la dificultad que supone completar con éxito los programas de doctorado: el tiempo medio necesario para la obtención de título de doctor¹ y las tasas de doctorados entre los estudiantes matriculados en el tercer ciclo². Al margen del coste personal que puede llegar a representar la barrera de la tesis doctoral para algunas personas, y de las repercusiones sobre la reputación de las universidades con bajos porcentajes de éxito (Pyke y Sheridan, 1993), es preciso considerar también la pérdida de varios años de actividad productiva de un número nada despreciable de los sujetos más inteligentes y creativos de cada generación (Gillingham, Seneca y Taussig, 1991), empleados en culminar un proyecto que, demasiadas veces, no se concluye³.

1 Varios estudios, llevados a cabo a finales de los 80, evidenciaban una clara tendencia al incremento del tiempo medio necesario para completar el doctorado a partir de los años 70. Ver al respecto Bowen y Sosa (1989), Duggan (1989), Filteau (1992), Moses (1985), Pyke y Sheridan (1993), Tuckman, Coyle y Bae (1990).

2 Existe un alto porcentaje de alumnos —alrededor del 85%, en el Estado Español— que no llegan a presentar la tesis doctoral, a pesar de haber seguido los cursos de doctorado. El problema alcanza tal magnitud que es designado como el Síndrome ABD (*All but dissertation*, o TMT —todo menos tesis—, en castellano).

3 Entre las consecuencias derivadas de no alcanzar el título de doctor, Pyke y Sheridan (1993) destacan, a nivel individual, la pérdida de tiempo y dinero, frustración de las ambiciones personales, metas no alcanzadas, rechazo social y sentimientos de incompetencia, desmoralización y depresión. A nivel institucional, señalan que las altas cifras de abandono pueden ser percibidas por los estudiantes y por la sociedad como reflejo de la incapacidad de la universidad para satisfacer las necesidades de sus alumnos y, en consecuencia, de una baja eficiencia.

La relación existente entre distintos tipos de disciplinas científicas y el tiempo necesario para completar el doctorado es, probablemente, uno de los datos más consistentes en la literatura sobre el tema (Baird, 1990, Bowen y Rundenstine, 1992, Duggan, 1989, Fletcher y Stren, 1992, Girves y Wemmerus, 1988, Hamilton, 1993, Holdaway, Deblois y Winchester, 1995, Moses, 1985, Sheridan y Pyke, 1994, Wright y Lodwick, 1989), llegando a explicar hasta un 26% de la varianza del éxito en los estudios de doctorado (Sheridan y Pike, 1994).

Dado que el problema presenta una especial incidencia en las áreas de ciencias sociales y humanidades, siendo mucho menor en las de ciencias biomédicas, ciencias experimentales y carreras técnicas, nuestro objetivo en este trabajo es el de analizar las diferencias existentes entre distintas disciplinas en relación con algunos factores determinantes del éxito en los estudios de tercer ciclo⁴.

No se defiende en este trabajo la conveniencia de equiparar o uniformar unas disciplinas con otras. La existencia de tasas más bajas de doctorado en ciencias sociales y humanidades, así como la mayor duración del período de realización de la tesis, no debería considerarse intrínsecamente negativo si los factores que actúan sobre las diferencias entre áreas están plenamente justificados y consensuados. Se pretende, por el contrario, detectar cuáles son esos factores en nuestro contexto educativo, analizar su naturaleza y pertinencia, con el fin de estudiar posibles propuestas de actuación para incrementar los porcentajes de éxito (mayor orientación al estudiante, programas de entrenamiento en supervisión de tesis, fomento de la financiación de proyectos de investigación en determinadas áreas, etc.).

Nos proponemos, por consiguiente, abordar, en primer lugar, el estudio sistemático de las variables que facilitan o dificultan la finalización de los estudios de doctorado, poniendo de relieve la influencia de tres grandes grupos de factores explicativos del éxito en alcanzar el título: económicos, personales y académicos. En la segunda parte del trabajo se analizan las diferencias entre disciplinas científico-tecnológicas versus humanísticas y sociales en relación con los tres grupos de variables previamente considerados.

Determinantes socioeconómicos

La utilización de modelos económicos para explicar las causas del progreso en los estudios de doctorado presupone un comportamiento racional en la toma de decisiones educativas individuales, en un contexto con limitaciones de recursos (de tiempo y dinero) a las que se ven confrontados los estudiantes. En estos modelos, la ayuda financiera recibida y la disponibilidad de tiempo para completar el programa serían dos condicionantes fundamentales del éxito (OCDE, 1987).

4 La definición de éxito que se considera en este trabajo se refiere a la consecución del título de doctor. No se olvida, sin embargo, que, para muchos estudiantes, el doctorado constituye una oportunidad asequible de formación continua, con lo que la apreciación subjetiva de éxito para estos alumnos no se vincula directamente a la consecución del título sino a la adecuación entre formación recibida y expectativas personales.

De acuerdo con este planteamiento, Gillingham, Seneca y Taussig (1991) consideran que la decisión de dejar o continuar el doctorado sería una consecuencia del análisis de los costes y beneficios esperados por el sujeto. Para algunos estudiantes, los beneficios de acabar lo más pronto posible son grandes y altamente probables (incorporación al cuerpo docente, por ejemplo), y los costos relativamente bajos. Para otros —quienes no trabajan en la universidad, especialmente—, los beneficios son marginales comparados con los costes (escasa valoración del título de doctor fuera del ámbito universitario⁵).

Las variables habitualmente consideradas en este tipo de estudios suelen ser las características demográficas, el tiempo dedicado al programa de doctorado, el horario laboral (para quienes trabajan), los ingresos y la ayuda financiera.

La cantidad y clase de ayuda económica recibida por el estudiante durante el doctorado (becas de investigación o docencia, fundamentalmente) parece ser —junto con la disciplina de estudio— uno de los mejores predictores del éxito en la consecución del título de doctor, según los análisis realizados desde planteamientos económicos (Bowen y Rudenstine, 1992; Dolph, 1983; Fletcher y Stren 1992; Girves y Wemmerus, 1988, OCDE, 1987; Pyke y Sheridan, 1993; Sheridan y Pyke, 1994). Abedi y Benkin (1987), por ejemplo, constatan que los estudiantes cuyos ingresos dependen de un empleo necesitan dos años más de tiempo para completar el doctorado frente a aquéllos que disponen de otras fuentes de ingresos, sobre todo becas de investigación o docencia. Existe, en este punto, un cierto consenso acerca de la ventaja que representa disfrutar de una beca, siempre y cuando las obligaciones asociadas a dicha ayuda no dificulten la dedicación al programa de doctorado, en cuyo caso puede convertirse en un inconveniente que alarga el plazo para la consecución del título (Sheridan y Pyke, 1994). Continúa siendo una cuestión de debate, sin embargo, si las tareas docentes encargadas a los becarios de doctorado perjudican la pronta realización de la tesis doctoral (Stirati y Cesaratto, 1995).

Cabe buscar la explicación a estos resultados en las mayores oportunidades de que gozan los becarios para implicarse activamente en los programas de doctorado, al verse libres de los contratiempos y preocupaciones que suele llevar aparejados un trabajo fuera de la universidad (Girves y Wemmerus, 1988). Tal como han señalado Gillingham, Seneca y Taussig (1991), las horas dedicadas a un empleo tienen un efecto negativo sobre el progreso en el doctorado, en la medida que impiden una mayor dedicación al programa. En correspondencia, los análisis de regresión efectuados en su estudio muestran también un mayor progreso entre quienes dedican un mayor número

5 No parece existir una gran diferencia entre licenciados y doctores en relación con las oportunidades de empleo en el mercado de trabajo no universitario. Los empresarios no suelen distinguir entre ambos títulos, tal como han constatado Pearson, Seccombe, Pyke y Connor (1993) en el ámbito de las ciencias sociales. En algunas especialidades, como educación, los estudiantes constatan que, una vez acabado el tercer ciclo, el hecho de ser doctores no les ha ayudado demasiado en sus carreras profesionales (Towsend y Mason, 1990). Aún cuando, en algunos casos, las posibilidades laborales son mejores para los doctores que para los licenciados, ello podría explicarse en términos de su mayor experiencia en el mercado de trabajo, puesto que buena parte de los doctores continúan realizando la misma actividad laboral que llevaban a cabo antes de obtener el título (Stirati y Cesaratto, 1995).

ro de horas al doctorado. En la misma línea, Sheridan y Pyke (1994) observan mayor rapidez en la obtención del título de doctor entre los estudiantes matriculados a tiempo completo en los programas de doctorado, frente a los alumnos matriculados a tiempo parcial.

Por otro lado, los becarios disponen también de mayores oportunidades para integrarse en los equipos de investigación y establecer relaciones con los profesores del departamento, lo cual estimula su motivación y compromiso para obtener el título de doctor; sin olvidar el beneficio percibido de mayores posibilidades de acceso a la carrera académica.

Determinantes personales

Entre las características individuales que suelen citarse como causa del progreso en los estudios de tercer ciclo destacan: la motivación individual, las habilidades y los rasgos de personalidad del estudiante (Willcoxson, 1994).

La motivación del alumno es argumentada como variable explicativa fundamental por la práctica totalidad de los profesores universitarios encuestados en distintos estudios (Hockey, 1996a, Holdaway, 1994). Como sea que la motivación (entusiasmo, perseverancia, curiosidad, etc.) es el constructo resultante de la interacción del resto de variables personales y contextuales, no abundaremos ahora en este tema.

Por lo que se refiere a las competencias y habilidades personales que debería poseer un buen estudiante de doctorado y un buen tesista, la principal fuente de los datos empíricos disponibles proviene de cuestionarios aplicados a profesores, directores de tesis, jefes de departamento o a los propios alumnos de doctorado (Hill, Acker y Black, 1994, Hockey, 1994, Parry, Atkinson y Delamont, 1994, Valarino, Meneses, Yáber y Pujol, 1996), y son resumidas por Hockey (1996a) de la siguiente manera: inteligencia, independencia, iniciativa, entusiasmo, creatividad, perseverancia y disponibilidad a dejarse aconsejar (p. 360), y por Moses (1985): diligencia, dedicación, trabajo duro, energía, entusiasmo, tenacidad sentido de la urgencia (p. 37); (ver también Valarino, 1992, para una completa revisión sobre el tema en castellano).

En una encuesta aplicada a 171 estudiantes de postgrado de la Universidad de Sidney, se manifestaban como dificultades más frecuentes, en relación con las habilidades necesarias para el progreso académico en los programas de tercer ciclo, la organización efectiva del tiempo (un 49% de los alumnos reconocían esta dificultad), la elaboración de un plan global de investigación y de un calendario de trabajo (42%), la clarificación de expectativas acerca del proceso de supervisión y la relación con el tutor (37%), selección y clarificación del tema de investigación (30,5%) y la redacción clara, lógica y coherente del trabajo de investigación (29%) (Willcoxson, 1994). Existen en la actualidad numerosos programas de ayuda dirigidos a mejorar las habilidades del alumno para superar las dificultades asociadas a la realización de la tesis.

Entre los rasgos de personalidad con mayor incidencia sobre la realización de la tesis doctoral merece la pena destacar la perseverancia, resistencia y optimismo, como condición necesaria para afrontar con éxito el aislamiento social que supone el proceso de investigación (Hockey, 1991), los sentimientos de desánimo ante las dificultades

experimentadas (selección del tema, clarificación de metas, superación del mito de la tesis perfecta y de la presión del tiempo, etc.), los sentimientos acerca de sí mismo (sensación de incompetencia que suele aparecer en momentos críticos, o cuando se compara la propia producción con la de otros compañeros o profesores, provocando ansiedad y bloqueo), y el manejo de las relaciones interpersonales durante el proceso (Valarino, 1992). La implantación de estilos de funcionamiento cooperativos en los departamentos universitarios parece producir beneficios de carácter personal, ya que permiten eludir el aislamiento, fomentar un sentido de pertenencia, bienestar psicológico, autorrealización, asunción de determinados riesgos y desarrollo de relaciones interpersonales (Anderson, 1996).

Determinantes académicos

Existen numerosas variables de carácter académico que inciden directamente sobre el éxito en la finalización del tercer ciclo universitario. En opinión de Hamilton (1993) la relación entre estudiante y tutor, la relación de los cursos de doctorado con la tesis, la excesiva ambición de los temas de investigación seleccionados y de los objetivos propuestos por el estudiante, junto a unas excesivas demandas de los directores, constituyen algunos de los condicionantes principales en los resultados finales.

Actualmente se dispone de evidencia suficiente acerca del proceso de supervisión de la tesis como variable explicativa fundamental del progreso en el doctorado (Anderson, 1996; Donald, Saroyan y Denison, 1995; Hockey, 1996a y 1996 b; Holdaway, 1994; Holdaway, Deblois y Winchester, 1995; Valarino, 1992; Willcoxson, 1994). Se ha demostrado un mayor progreso en la planificación del trabajo de investigación al final del primer año en el programa de doctorado entre aquellos estudiantes que mantienen un contacto más frecuente con su director (Wright y Lodwick, 1989). Las variables referidas al rol del tutor, sus características personales-emocionales y sus competencias (Valarino, Meneses, Yáber y Pujol, 1996), y más específicamente su conocimiento del campo de investigación y su disponibilidad (Donald, Soroyan y Denison, 1995), así como la calidad de las relaciones interpersonales estudiante-tutor (Ballard y Clanchy, 1993; Salmon, 1992) se han mostrado especialmente relevantes.

Igualmente, como se ha señalado en la introducción, los estudios descriptivos muestran un acuerdo unánime en señalar importantes diferencias entre distintas tipologías de áreas científicas en cuanto al ritmo y resultados alcanzados por los estudiantes en los programas de doctorado. A pesar de las dificultades para comparar datos obtenidos con distintos criterios metodológicos, en distintos momentos y en contextos diversos, algunas investigaciones reflejan diferencias entre uno (Gillingham, Seneca y Taussig, 1991; Yeates, 1991) y tres años (Duggan, 1989) de tiempo medio requerido para obtener el título de doctor, en favor de los estudiantes de disciplinas científicas y técnicas frente a los de ciencias sociales y humanidades.

Se ha propuesto, como variable explicativa de este dato, las dificultades de carácter económico que afrontan los alumnos de doctorado y que se concentran principalmente en las áreas de humanidades y de ciencias sociales (Fletcher y Stren, 1992), probablemente a causa de una menor disponibilidad de becas de investigación en estas

áreas, de una mayor afluencia de alumnos y de la necesidad de muchos de ellos de mantener un empleo paralelamente a la realización del doctorado.

Una segunda línea argumental se refiere a los estilos de trabajo y supervisión particulares de cada área científica y a la naturaleza y papel que desempeña la tesis doctoral en cada una de ellas o entre grupos de disciplinas.

Algunos estudios llevados a cabo en universidades canadienses ponen en evidencia notables diferencias en las prácticas de supervisión llevadas a cabo en distintos ámbitos científicos (Donald, Saroyan y Denison, 1994; Holdaway, Deblois y Winchester, 1995), siendo generalmente más efectivas en las áreas de ciencias experimentales, ciencias de la vida y carreras técnicas (por ejemplo, el grado de vinculación de la tesis con la línea de investigación del tutor, la ayuda en la preparación de publicaciones, la motivación y orientación continuada del estudiante, entre otros aspectos, es mayor en estas áreas). Wright y Lodwick (1989) también constatan una menor frecuencia de supervisión del trabajo del estudiante por parte del director en el área de ciencias sociales, siendo estos estudiantes los más lentos en avanzar en la planificación de su trabajo durante el primer año y los más pesimistas con respecto a las previsiones de acabar la tesis en los plazos fijados. En general, se espera que los estudiantes de ciencias sociales planteen sus propios temas de investigación, elaboren su propio plan de trabajo y manifiesten un alto grado de independencia y autonomía respecto al tutor (Anderson, 1996; Hockey, 1996b; Moses, 1992; O'Brien, 1995). De hecho, los datos disponibles confirman que la elección del tema de investigación por parte del estudiante es más frecuente en educación, ciencias sociales y humanidades que en el resto de disciplinas científicas, en las que el director juega un papel más relevante en la decisión final, proponiendo el tema en muchos casos (Holdaway, Deblois y Winchester, 1994). Esta «tradicción» de mayor autonomía del estudiante se refleja, por ejemplo, en el número de publicaciones conjuntas derivadas de la investigación doctoral. Según datos aportados por Anderson (1996), la coautoría de las publicaciones entre supervisores y estudiantes es más frecuente en ciencias experimentales. En la misma línea, Holdaway, Deblois y Winchester (1995) constatan una mayor frecuencia de publicaciones firmadas por estudiantes de doctorado como autor único en el área de humanidades, frente al resto de disciplinas.

Smart, Baird y Bode (1991) y Smart y Hagedorn (1994), por su parte, señalan diferencias en cuanto a las estrategias educacionales utilizadas en la enseñanza de postgrado y las demandas de aprendizaje en distintas disciplinas, denunciando, al mismo tiempo, la ausencia prácticamente total de investigación sobre el tema.

Hamilton (1993) destaca el hecho de que en ciencias sociales y humanidades parece existir una cierta desconexión entre el contenido de los cursos de doctorado y la tesis doctoral. Es habitual que los cursos se centren en aspectos generales de un área de investigación, o en sus enfoques metodológicos, mientras que en el caso de un biólogo molecular, por ejemplo, los contenidos y actividades de los cursos de doctorado están subordinados a la participación del alumno en las tareas del laboratorio, donde, trabajando codo a codo con su director, toma parte activa en la investigación. Por otra parte, es costumbre que los estudiantes de las áreas científico-tecnológicas seleccionen el tema de tesis durante el primer año del programa de doctorado, iniciando la

investigación de manera inmediata y paralela a la realización de los cursos. En contraposición, los estudiantes de ciencias sociales y humanidades no suelen pensar seriamente en la tesis hasta que no han acabado los cursos de doctorado. De alguna manera, considera Hamilton, en estas disciplinas la tesis es a menudo tangencial al proceso de formación que representan en sí mismos los cursos de doctorado. Por otro lado, quizá los objetivos que plantean suelen ser demasiado ambiciosos, su extensión excesiva y exagerado el tiempo necesario para ser completadas con éxito (Becher et al., 1994).

La conclusión que se desprende de la revisión de la literatura que hemos llevado a cabo puede resumirse en la escasez de datos substanciales referidos a la educación de postgrado (Baird, 1990), la ausencia de un marco conceptual global, a pesar de algunos esfuerzos aislados (Holdaway, 1994) y de una cierta dispersión metodológica. Queda sin explorar, además, la relación entre enseñanza y aprendizaje, que no parece preocupar, por el momento, a los investigadores en este campo, más centrados en explicar las tasas de éxito-fracaso que en el análisis de los procesos educativos (Smart y Hagedorn, 1994).

El objetivo principal de este trabajo es el de contribuir al análisis de las posibles diferencias existentes entre diversas áreas disciplinares en relación con los tres grupos de factores determinantes del éxito en el doctorado, a partir del estudio de un caso concreto. Con este propósito se exponen los resultados de una investigación de carácter descriptivo llevada a cabo en la Universidad de Girona (UdG).

METODOLOGÍA

Descripción de la muestra

Han participado en la investigación 81 estudiantes (34,03%), matriculados en el primer año de los programas de doctorado correspondientes al bienio 1995-97 de la UdG, sobre un total de 238 estudiantes matriculados en dicho bienio⁶.

Por sexo, 50 son hombres (61,73%) y 31 mujeres (38,27%), porcentajes que no difieren significativamente ($\chi^2_1 = 2,28$ $p = n.s.$) de la población de primer año de la UdG (53,36% hombres y 46,64% mujeres). Existen diferencias significativas ($\chi^2_1 = 4,21$ $p = 0,04$) con relación a los porcentajes encontrados en el conjunto de la población española matriculada en los programas de doctorado para ese mismo curso 1995-96 (50,33% hombres y 49,67% mujeres).

Si se considera la disciplina de estudio, la muestra está formada por 39 estudiantes de humanidades y ciencias sociales (48,15%)⁷ y 42 estudiantes de programas científico-tecnológicos (51,85%)⁸, porcentajes que difieren significativamente ($\chi^2_1 = 4,81$ $p = 0,03$)

6 La complejidad organizativa de los estudios de doctorado, la dispersión y dificultad de localización de los estudiantes en las aulas, junto con la excesiva longitud del cuestionario utilizado para la recogida de información, son la causa del bajo porcentaje de participación.

7 Incluye los doctorados de Derecho privado, Lengua y literatura, Historia, Filosofía y Psicología.

8 Incluye los doctorados de Microbiología y bioquímica aplicadas, Física, Biología ambiental, Química computacional, Química, Organización y gestión de sistemas productivos y Tecnologías del medio ambiente.

de los observados en la población de primer año de la UdG (60,08% ciencias sociales y humanidades y 39,92% en ciencias). Las proporciones de hombres y mujeres en cada área disciplinar son similares al total de la población de primer año de doctorado de la UdG ($\chi^2_1=0,15$ en letras y $\chi^2_1=2,99$ en ciencias $p = n.s.$).

Instrumento de recogida de datos

El cuestionario aplicado al final del primer año del programa consta de 87 preguntas referidas a cuatro grandes bloques: datos académicos, sociodemográficos y laborales; expectativas y motivaciones para realizar los estudios de tercer ciclo; opiniones sobre dichos estudios y estado actual y previsión de tiempo respecto a la tesis doctoral. Se consideraran, a efectos de comparación entre disciplinas, las siguientes variables:

a) *Factores socioeconómicos*: edad; situación laboral; salario; horario laboral (horas/semana); estabilidad laboral percibida (las tres últimas variables, sólo para los estudiantes ocupados); cargas familiares (número de hijos); fuente principal de ingresos; vinculación docente a la universidad; y disfrute de becas de docencia — investigación.

b) *Factores personales*: Por lo que se refiere a las motivaciones que han influido en la decisión de estudiar tercer ciclo se han analizado las razones por las cuales decidieron realizar el doctorado (8 preguntas con 6 opciones de respuesta en una escala de 0= 'este factor no ha tenido ninguna importancia en el momento de tomar la decisión' a 5= 'este factor ha tenido una gran importancia'); objetivos personales a los que responde la formación de doctorado que están cursando (9 preguntas valoradas de 0= 'no es un objetivo prioritario' a 5= 'es un objetivo prioritario'); expectativas acerca de los resultados que pueda comportar la realización del doctorado y la valoración subjetiva de dichos resultados (6 preguntas, de 0= 'resultado nada probable' y 'resultado que no tiene ningún valor' a 5= 'resultado muy probable' y 'resultado muy importante'); y dificultades o inconvenientes que supone cursar los estudios de tercer ciclo (7 preguntas, de 0= 'no supone ningún inconveniente para mí' a 5= 'constituye un inconveniente muy importante').

En relación con las habilidades o capacidades percibidas para la realización de la tesis doctoral, se les pidió una valoración —de 0 a 5— de su situación personal en relación con una serie de competencias o características que podrían influir en su elaboración (conocimientos metodológicos, capacidad de trabajo, etc.).

c) *Factores académicos*. Se han tenido en cuenta cuatro grupos de variables:

Se han evaluado, por un lado, las variables directamente relacionadas con los cursos de doctorado: Dificultad percibida de los cursos (alta, media, baja); percepción subjetiva del esfuerzo invertido (alto, medio, bajo); grado de cumplimiento en la asistencia a clase y realización de las tareas asignadas en el programa (alto, medio, bajo); número de horas por semana que dedican a asistir a las clases de doctorado, y al estudio y trabajo personal en casa; características o requisitos que deberían tener los cursos de doctorado (11 ítems que responden a la pregunta: 'si tu pudieras diseñar el programa de tercer ciclo, ¿qué requisitos incluirías en su diseño?'); satisfacción con diversos

aspectos del programa (7 preguntas acerca de si se han cumplido las expectativas de entrada con respecto a la metodología docente, profesorado, infraestructuras, etc. y 2 preguntas en las que se les pide que indiquen si recomendarían los estudios de tercer ciclo que están cursando a otras personas y si volverían a matricularse en el programa en caso de empezar nuevamente).

Un segundo grupo de preguntas se relaciona con el grado de planificación del trabajo de doctorado y de la tesis doctoral (si piensan llevarlos a cabo, si han elegido tema y director y si han elaborado un plan de trabajo), y 2 preguntas acerca del tiempo previsto para completar los cursos de doctorado y la tesis doctoral.

Por otro lado, teniendo en cuenta la posible influencia de los conocimientos y aptitudes personales en la realización de la tesis doctoral, se les ha pedido que evalúen la formación académica recibida durante la licenciatura (nada o poco satisfactoria, más o menos satisfactoria y bastante o muy satisfactoria) y la nota promedio de la carrera.

Se ha evaluado también la productividad científica relacionada con los estudios de doctorado a través de cinco preguntas sobre el número de publicaciones, en solitario o en colaboración, vinculadas a los estudios de tercer ciclo (comunicaciones, colaboraciones en la prensa escrita, artículos especializados, capítulos de libro y libros).

RESULTADOS

Se presentan a continuación las frecuencias y porcentajes⁹ obtenidos por los dos grupos de disciplinas considerados en la investigación, para cada alternativa de respuesta, en cada uno de los grupos de variables considerados en el estudio.

En la tabla 1 se muestran los resultados correspondientes al análisis de algunos factores socioeconómicos susceptibles de condicionar el progreso durante los estudios de doctorado.

En general, el perfil de los estudiantes de doctorado que han participado en la investigación es el de un colectivo que se mantiene principalmente con sus propios ingresos, aun cuando un 25,35% se encuentra en situación de paro y sólo un 42% puede considerarse totalmente independiente desde el punto de vista económico. Esta situación afecta por igual a los estudiantes de ciencias y a los de humanidades y ciencias sociales.

Las únicas variables de carácter económico respecto a las cuales ambos grupos de disciplinas se comportan de manera distinta son la vinculación a la universidad, mayor entre los estudiantes de ciencias y carreras técnicas (69%) y, relacionada con ella, el disfrute de becas de docencia o investigación, mucho mayor también en este grupo (78,3%). No se aprecian diferencias importantes en ninguno de los otros factores considerados.

⁹ Todos los porcentajes de las tablas se han calculado sobre el total de estudiantes de letras y ciencias que contestan a una misma alternativa de respuesta. Se trata, por tanto, de porcentajes referidos a cada grupo disciplinar con respecto al otro.

Tabla 1
FACTORES SOCIOECONÓMICOS

Variable * Alternativa de respuesta	Disciplina [n (%)]		Contraste
	Hum.-CC. sociales	CC exper.-tecnol.	
<i>Situación laboral</i> 1. Ocupado 2. Parado	26 (49,1) 9 (50,0)	27 (50,9) 9 (50,0)	$\chi^2_1 = 0,005$ p= n.s.
<i>Salario (sólo ocupados)</i> 1. < 100.000 ptas./mes 2. 100.000-175.000 ptas./mes 3. > 175.000 ptas./mes	6 (50,0) 10 (45,5) 9 (60,0)	6 (50,0) 12 (54,5) 6 (40,0)	$\chi^2_2 = 0,76$ p= n.s.
<i>Estabilidad laboral percibida (sólo ocupados)</i> 1. Poco o nada estable 2. Más o menos estable 3. Bastante o totalmente estable	7 (50,0) 9 (56,25) 12 (54,55)	7 (50,0) 7 (43,75) 10 (45,45)	$\chi^2_2 = 0,12$ p= n.s.
<i>Forma de subsistencia</i> 1. Vive exclusivamente de sus propios ingresos 2. Ingresos propios con alguna ayuda de la familia 3. Recursos familiares y algún ingreso propio 4. Exclusivamente de los recursos familiares	19 (55,9) 9 (42,9) 9 (37,5) 2 (100)	15 (44,1) 12 (57,1) 15 (62,5) —	$\chi^2_3 = 4,29$ p= n.s.
<i>Beca</i> 1. Sí 2. No	5 (21,7) 30 (63,8)	18 (78,3) 17 (36,2)	$\chi^2_1 = 10,94$ p= 0,0009
<i>Vinculación profesional a la universidad</i> 1. Sí 2. No	9 (31,0) 30 (58,8)	20 (69,0) 21 (41,2)	$\chi^2_1 = 5,71$ p= 0,017
<i>Horario laboral (sólo ocupados)</i>	media = 34,07 SD=10,15 (n=23)	media = 36,62 SD=11,35 (n=21)	$t_{42} = 0,79$ p= n.s.
<i>Cargas familiares (número de hijos)</i>	media = 0,26 SD=0,55 (n=39)	media = 0,14 SD=0,47 (n=42)	$t_{79} = 1,00$ p= n.s.
<i>Edad (años)</i>	media = 29,36 SD=6,00 (n=39)	media = 27,67 SD=4,62 (n=42)	$t_{79} = 1,41$ p= n.s.

En lo que respecta a los factores de tipo motivacional, no se han encontrado diferencias significativas en ninguna de las variables relacionadas con las expectativas de resultados, ni con la valoración de dichos resultados ('el título de doctorado me permitirá encontrar un trabajo más satisfactorio que el actual', 'me permitirá encontrar un trabajo', 'un título de mayor nivel académico', 'mayor número de conocimientos de mi profesión', 'conocer a otras personas y hacer nuevos contactos o amistades', 'mantenerme ocupado'). En todos los casos, las medias de las puntuaciones obtenidas en una escala de 0 a 5 oscilan entre 1,46 y 2,70, no encontrándose diferencias estadísticamente significativas entre los dos grupos considerados en el estudio (valor máximo de $t < 1,56$ en cada una de las 12 preguntas, $p = \text{n.s.}$).

Tampoco se han encontrado diferencias significativas entre disciplinas en lo que se refiere a las dificultades o inconvenientes que supone la realización del doctorado ('tiempo que podrían dedicar a otras cosas', 'sobrecarga de la vida cotidiana', 'desplazamientos', 'coste económico de los estudios' y 'falta de tiempo para las relaciones sociales'), que podrían suponer algún tipo de desmotivación para la finalización de la tesis. En general, los estudiantes de doctorado de la muestra consideran que se trata de inconvenientes poco importantes (todas las medias de ambos grupos son inferiores a 2,05; valor máximo de $t < 1,62$, $p = \text{n.s.}$). Únicamente aparecen diferencias en lo que respecta al hecho de que 'la familia no puede gozar de su tiempo libre' (los estudiantes de ciencias consideran este inconveniente de mayor importancia: media= 1,60 SD= 0,80) que los de letras (media= 1,26 SD= 0,50; $t_{78} = 2,20$ $p = 0,028$).

Un comportamiento distinto muestran ambos grupos de disciplinas con relación a las razones y objetivos personales que han llevado a los estudiantes a cursar el doctorado. En la tabla 2 se recogen los resultados obtenidos para estas variables.

De acuerdo con los resultados, los estudiantes de letras parecen haber tenido en cuenta —en mayor medida que los de ciencias— el prestigio social y profesional que puede proporcionar el título de doctor, la necesidad de una mayor formación académica y el hecho de no desvincularse totalmente de la universidad. Ello puede explicarse teniendo en cuenta la mayor presencia de becarios en ciencias, hecho que les garantiza la vinculación universitaria.

El análisis de los objetivos personales a los que obedece la formación de tercer ciclo confirma también las diferencias halladas en las razones aducidas para la realización del doctorado. Nuevamente, se observa un comportamiento distinto entre ambas áreas disciplinares en lo referente a la búsqueda de una mayor formación académica y de un mayor estatus —a través del prestigio que puede proporcionar el título de doctor entre los estudiantes de letras—.

En las preguntas referidas a su situación personal con respecto a la tesis doctoral, y a la evaluación de las propias competencias para llevarla a cabo con éxito, los estudiantes de ambos grupos coinciden en valorar de forma parecida sus conocimientos metodológicos, teóricos, capacidad de trabajo, fuerza de voluntad, apoyo familiar, tiempo disponible y haber elegido un buen tema de trabajo. Sólo se han encontrado diferencias significativas entre ambos grupos en lo que se refiere a la elección del director; los estudiantes de ciencias consideran que han elegido un buen director en mayor medida que los estudiantes de letras (medias= 4,50 —SD= 0,67— y 3,65 —SD= 1,76—

Tabla 2

RAZONES Y OBJETIVOS PERSONALES PARA LA REALIZACIÓN DEL DOCTORADO

Variable	Disciplina [media (SD)]		t	Significación
	H-CC. sociales	CC. exper.-Tecn.		
<i>Razones</i>	(n= 39)	(n= 41)		
Mantenerse al día en el trabajo actual	1,95 (0,86)	1,61 (0,83)	1,79	p= n.s.
Prestigio personal/profesional	1,92 (0,87)	1,55 (0,77)	2,06	p= 0,043
Obtener un currículum diferenciador de otros titulados	1,82 (0,79)	2,23 (0,82)	2,33	p= 0,022
Especialización para superar déficits de la licenciatura	1,90 (0,79)	2,12 (0,77)	1,28	p= n.s.
Mayor formación académica	2,72 (0,46)	2,40 (0,70)	2,40	p= 0,019
Acceso a una plaza docente en la universidad	2,13 (0,86)	1,83 (0,80)	1,60	p= n.s.
Mantener un mayor contacto con la universidad	2,36 (0,63)	1,93 (0,81)	2,66	p= 0,009
Perfeccionamiento y reciclaje profesional	1,85 (0,78)	1,78 (0,86)	0,38	p= n.s.
<i>Objetivos</i>	(n= 39)	(n= 41)		
Formación investigadora	2,56 (0,68)	2,57 (0,70)	0,05	p= n.s.
Formación profesional de carácter práctico	1,95 (0,65)	2,05 (0,85)	0,59	p= n.s.
Formación profesional de carácter teórico	2,36 (0,67)	2,02 (0,78)	2,07	p= 0,042
Formación profesional de carácter metodológico	2,51 (0,68)	2,07 (0,81)	2,64	p= 0,010
Formación docente	1,97 (0,81)	1,52 (0,77)	2,56	p= 0,012
Formación cultural general	2,00 (0,76)	1,79 (0,78)	1,25	p= n.s.
Promoción laboral	1,56 (0,72)	1,79 (0,90)	1,23	p= n.s.
Promoción personal (estatus o prestigio)	1,97 (0,75)	1,64 (0,73)	2,00	p= 0,049
Promoción académica (titulación superior a la actual)	2,18 (0,76)	2,37 (0,77)	1,09	p= n.s.

respectivamente; $t_{62} = 2,53$ $p = 0,013$). La formulación de la pregunta no permite saber si los estudiantes de humanidades y CC sociales consideran que su director no es tan bueno o, simplemente, es que no han buscado director todavía. Más adelante se aclarará esta cuestión.

Para el análisis de los factores académicos dividiremos la exposición en cinco apartados: planificación del proyecto de investigación y tesis doctoral, dedicación a los estudios, formación universitaria, requisitos formativos exigidos a los cursos y satisfacción con los mismos, y productividad científica.

Los resultados obtenidos en relación con el nivel de planificación del trabajo de investigación y de la tesis doctoral se muestran en la tabla 3. Todos, excepto un estudiante de ciencias, piensan llevar a cabo el trabajo de investigación, mientras que sólo 2 estudiantes de letras y 4 de ciencias no piensan realizar la tesis doctoral (9,23%), sobre 65 que contestan la pregunta.

De los datos de la tabla se deduce que los estudiantes del área científico-tecnológica empiezan a trabajar antes en su proyecto de investigación que sus compañeros de

Tabla 3

NIVEL DE PLANIFICACIÓN DEL TRABAJO DE INVESTIGACIÓN Y DE LA TESIS DOCTORAL AL FINAL DEL PRIMER AÑO EN EL PROGRAMA

Variable (alternativas de respuesta: si/no)	Disciplina		Contraste	Significación
	Hum.-CCsoc. [n (%)]	CCexp.-tecn.		
Han pensado el tema del trabajo de investigación	25 (47,2)/8 (61,5)	28 (52,8)/5 (38,5)	$\chi^2_1 = 0,86$	p= n.s.
Han elegido el director	18 (40,0)/16 (69,6)	27 (60,0)/7 (30,4)	$\chi^2_1 = 5,32$	p= 0,021
Han elaborado un plan de trabajo	14 (35,9)/19 (73,1)	25 (64,1)/7 (26,9)	$\chi^2_1 = 8,63$	p= 0,003
Han pensado el tema del trabajo de la tesis doctoral	20 (45,5)/12 (63,2)	24 (54,5)/7 (36,8)	$\chi^2_1 = 1,66$	p= n.s.
Han elegido el director	15 (38,5)/18 (72,0)	24 (61,5)/7 (28,0)	$\chi^2_1 = 6,86$	p= 0,008
Han elaborado un plan de trabajo	9 (29,0)/24 (75,0)	22 (71,0)/8 (25,0)	$\chi^2_1 = 13,34$	p= 0,0002
Previsión de tiempo para acabar los cursos de doctorado (en años)	media=2,00 (SD=0,08) n=32	media=2,40 (SD=1,32) n=40	t= 1,59	p= n.s.
Previsión de tiempo para acabar la tesis doctoral (en años adicionales)	media=3,48 (SD=1,79) n=29	media=2,73 (SD=1,35) n=33	t= 1,89	p= 0,063 n.s.

humanidades y ciencias sociales. De hecho, el 79,41% de los 34 estudiantes de ciencias experimentales y tecnologías ya han elegido un director para el trabajo de doctorado al final del primer año y el 78,13% han elaborado un plan de trabajo, frente al 52,94% y 42,42% respectivamente en ciencias sociales y humanidades. De la misma forma, sólo un 27,27% de estos últimos han elaborado un plan de trabajo para la tesis doctoral, frente al 73,33% del área científico-tecnológica, mientras que un 45,45% han elegido director de tesis por un 77,42% en ciencias.

De acuerdo con esta dinámica, también se observa que el tiempo previsto para la finalización de la tesis es mayor en humanidades y ciencias sociales; las diferencias entre ambos grupos muestran una tendencia a la significación ($p= 0,063$). Los estudiantes de ciencias tienen previsto, por el contrario, tardar algo más en completar los cursos de doctorado, aunque no se observan diferencias estadísticamente significativas entre disciplinas.

Con respecto a la dedicación en los estudios de doctorado, ambos grupos muestran un alto grado de cumplimiento en cuanto a la asistencia a clase y a la realización de las tareas asignadas por los profesores. Ningún estudiante reconoce un nivel de cumplimiento bajo. El 64,86% de estudiantes en humanidades y ciencias sociales y el 83,33% en ciencias informan de una alta dedicación. Se observa un cierto grado de asociación entre ambas variables (dedicación y tipo de disciplina), aunque no llega a ser estadísticamente significativa ($\chi^2_1 = 3,55$ $p= 0,059$). Por otra parte, la distribución de las variables 'percepción del esfuerzo dedicado a los estudios de doctorado' y 'dificultad percibida de los mismos' es prácticamente la misma en ambos grupos disciplinares ($\chi^2 = 3,89$ y $\chi^2 = 0,75$, $p=$ n.s. Mayoritariamente se dedica un esfuerzo moderado y la dificultad percibida es intermedia). No se observan tampoco diferencias

significativas en lo referente al número de horas de estudio o trabajo personal en casa, relacionado con el tercer ciclo ($t_{70} = 0,98$ $p = n.s.$), aunque es algo mayor en letras —10,89h/semana de promedio frente a 8,29 en ciencias, si bien existe una alta dispersión en ambas distribuciones (desviaciones estándar de 8,98 y 13,19 horas respectivamente—. En cambio, se observan diferencias estadísticamente significativas en lo referente al número de horas de asistencia a las clases programadas (5,28 h/semana de promedio en letras —SD= 1,65— y 7,76 —SD= 5,75— en ciencias; $t_{68} = 2,43$ $p = 0,020$).

Considerados en conjunto el trabajo personal y la asistencia a las clases programadas obtenemos un promedio de 15,77 horas semanales de dedicación a los estudios de doctorado (SD= 12,17), con escasas diferencias entre ciencias y letras (16,09 en ciencias —SD= 15,30— por 15,46 en letras —SD= 8,44—, $t_{68} = 0,21$ $p = n.s.$).

En lo que respecta a la formación previa del estudiante, ambos grupos presentan una distribución similar en sus respuestas sobre el grado de satisfacción con la formación recibida durante la licenciatura que, en general, se considera satisfactoria (sobre un total de 78 alumnos, el 92,11% en letras y el 77,5% en ciencias la consideran satisfactoria o muy satisfactoria). Por lo que se refiere a la nota promedio de licenciatura no se observan diferencias significativas entre los estudiantes del área de humanidades y ciencias sociales (promedio de 2,04 SD= 0,50) frente a los del área científico-tecnológica (1,95 SD= 0,70).

Relacionado con los aspectos formativos necesarios para la realización del doctorado los estudiantes reclaman de los cursos una serie de requisitos que se recogen en la tabla 4.

La relación entre los requisitos exigidos a los cursos y el tipo de disciplina académica únicamente se muestra significativa en lo que respecta a la exigencia de mayor número de prácticas profesionales en el área científico-tecnológica. En el resto de los aspectos considerados, ambos grupos distribuyen sus respuestas de forma similar,

Tabla 4

REQUISITOS QUE INCLUIRÍAN LOS CURSOS DE DOCTORADO CASO DE SER DISEÑADOS POR LOS PROPIOS ESTUDIANTES
(frecuencia (y %) de las respuestas afirmativas a cada pregunta)

Requisitos exigidos a los cursos	H-CCsoc. (n=38)	CCexp.-tec. (n=41)	χ^2_1	Significación
Contenidos más específicos	15 (50,0)	15 (50,0)	0,70	p= n.s.
Contenidos más generales	0	5 (100)	—	—
Más prácticas profesionales	10 (32,3)	21 (67,7)	5,13	p= 0,024
Más aspectos metodológicos	17 (47,2)	19 (52,8)	0,02	p= n.s.
Mayor incidencia en la investig. básica	20 (52,6)	18 (47,4)	0,60	p= n.s.
Más formación docente	11 (68,8)	5 (31,3)	3,43	p= 0,064 n.s.
Más formación investigadora	23 (48,9)	24 (51,1)	0,04	p= n.s.

aunque los de letras reclaman una mayor formación docente, aspecto que confirma las diferencias entre ambos grupos respecto a la formación docente como objetivo personal que justifica la realización del tercer ciclo (ver tabla 2).

En referencia a la satisfacción de los estudiantes con los cursos de doctorado cabe decir que debe aceptarse la hipótesis de independencia entre las variables que evalúan este aspecto y el tipo de disciplina académica. El comportamiento de ambos grupos en lo que se refiere a la satisfacción de sus expectativas con respecto a la metodología docente, calidad del profesorado, posibilidades de investigación, servicios a su disposición, y atención recibida, son similares. Únicamente aparecen diferencias entre disciplinas en lo que se refiere a la satisfacción con los contenidos del programa (el 68,42% de letras está muy satisfecho frente a un 15,79% satisfecho y un 15,79% nada satisfecho, mientras que los porcentajes en ciencias son del 41,46%, 34,15% y 24,39% respectivamente; $\chi^2= 5,98$ $p= 0,05$) y en la disponibilidad de infraestructuras, más satisfactoria en letras ($\chi^2= 6,14$ $p= 0,046$).

Los porcentajes de estudiantes que recomendarían los cursos a otras personas, y que volverían a matricularse de nuevo, son altos en ambos grupos (86,84% y 94,74% en letras, frente al 73,17% y 82,93% en ciencias).

Finalmente, con respecto al número de publicaciones, en solitario o en colaboración, vinculadas a los estudios de tercer ciclo, cabe señalar que la productividad es baja (no hay que olvidar que se trata del primer año de doctorado). Los promedios (y desviación estándar) en el área humanística y en el área científica son los siguientes: 0,06 (0,24) y 1,10 (1,50) para la presentación de comunicaciones en congresos; 0,15 (0,36) y 0,31 (1,23) de colaboraciones en la prensa escrita; 0,18 (0,53) y 0,41 (0,87) de artículos en revistas especializadas; 0 y 0,04 (0,19) de capítulos de libro y 0 en ambos casos para la publicación de libros. Sólo se observan diferencias estadísticamente significativas en lo que respecta a la presentación de comunicaciones en jornadas y congresos ($t_{60}= 3,71$ $p= 0,001$).

DISCUSIÓN

En conjunto, los resultados de la investigación son consistentes con los datos obtenidos en otros contextos educativos, en referencia a las diferencias observadas entre estudiantes de doctorado de disciplinas científico-tecnológicas y humanístico-sociales con respecto a determinados factores socioeconómicos, personales y académicos que condicionarían el progreso en los estudios de tercer ciclo. Estas diferencias se centran, básicamente, en dos aspectos: la vinculación del estudiante a la universidad y, probablemente relacionado con ello, el proceso de planificación y elaboración del trabajo de investigación y la tesis doctoral.

Nos referíamos en la introducción a la influencia que, según el planteamiento de los modelos económicos, ejercería la variable 'cantidad y tipo de ayuda económica recibida por el estudiante', sobre el progreso en el doctorado (principalmente el hecho de ser becario). Nuestros resultados confirman, efectivamente, la presencia de mayor número de becarios en ciencias, lo que hipotéticamente justificaría los mayores porcentajes de éxito en este grupo de disciplinas. La explicación económica a estos datos se apoya

en la posibilidad de una mayor dedicación de los becarios a los estudios de doctorado. Por nuestra parte, no observamos, sin embargo, grandes diferencias en cuanto al número de horas dedicadas al tercer ciclo entre los estudiantes de ciencias y los de letras. En cambio, las diferencias son muy notables en el nivel de planificación de los proyectos de investigación y en las previsiones de tiempo para acabar la tesis doctoral. En consecuencia, nuestra interpretación de la existencia de mayores tasas de éxito en ciencias apuntan en la línea de una mayor integración de los estudiantes en equipos o proyectos de investigación.

En este sentido, los resultados de Wright y Lodwick (1989) demostraban que el primer año es crítico para una pronta elaboración de la tesis doctoral ya que buena parte del trabajo preparatorio se hace entonces. Según estos autores, al final del primer año los alumnos tendrían un buen conocimiento del área temática de trabajo gracias a unas amplias lecturas, habrían decidido un problema de investigación y una metodología de trabajo, habrían establecido algún tipo de rutina para organizar el material y habrían descubierto dónde y cómo encontrar la información que necesitan y establecido una relación con el profesorado de su departamento, con sus supervisores y sus compañeros. No cabe duda de que todo ello es mucho más sencillo cuando se trabaja dentro de un equipo de investigación, con proyectos en marcha claramente definidos, tanto en los objetivos teóricos y funcionales como en los aspectos metodológicos y, a menudo, con un alto grado de concreción y aplicabilidad, características que suelen darse con mayor frecuencia en ciencias que en humanidades y en ciencias sociales.

En nuestro estudio, los alumnos del área científico-tecnológica presentan una clara ventaja en lo que se refiere a la elección del director y la elaboración de un plan de trabajo para la tesis doctoral. Serían necesarias, sin embargo, nuevas investigaciones que permitieran profundizar un poco más en la explicación de estos resultados, con el fin de determinar hasta qué punto obedecen a la existencia de prácticas habituales de colaboración científica en determinados campos, a la relación de estas prácticas con la intervención administrativa a través de subvenciones a proyectos presentados por equipos de investigación, o a formas de trabajo tradicionales asociadas a determinadas áreas disciplinares, como propone Hamilton (1993).

Nuestros resultados parecen confirmar al menos dos de las hipótesis planteadas por Hamilton: la primera, tal como señalábamos en la introducción, se refiere a la costumbre entre los estudiantes de las áreas científico-tecnológicas de seleccionar el tema de tesis durante el primer año del programa de doctorado, iniciando la investigación de manera inmediata y paralela a la realización de los cursos, mientras que los estudiantes de ciencias sociales y humanidades no suelen pensar seriamente en la tesis hasta que no han acabado los cursos de doctorado.

La segunda hipótesis planteada por Hamilton hace referencia a la posible desconexión entre el contenido de los cursos de doctorado y la tesis doctoral, siendo habitual en las disciplinas humanísticas y sociales que los cursos se centren en aspectos generales del área de conocimiento. Un repaso a los resultados presentados en la tabla 4 permite constatar una mayor demanda de contenidos generales en los cursos por parte de los alumnos del área de ciencias, mientras que ningún alumno de letras reclama este tipo de contenidos. También Holdaway (1994) encuentra diferencias entre las opi-

niones de los supervisores de distintas disciplinas acerca de los cursos de doctorado. El 56% de directores de tesis en ciencias biomédicas y el 66% en ingenierías y ciencias experimentales estarían de acuerdo en exigir la realización de los cursos a los alumnos, frente a un 89% en humanidades y un 77% en ciencias sociales, lo que constataría el mayor carácter formativo que desempeñarían en este último grupo de disciplinas.

Por lo demás, tanto unos como otros parecen desear más contenidos específicos, metodológicos, investigación básica y formación investigadora en proporciones similares y bastante altas. Ello nos lleva a plantear si realmente se consideran bien preparados para afrontar la realización de la tesis doctoral y si los cursos responden a sus necesidades reales. Los resultados encontrados por Stirati y Cesaratto (1995), en Italia, muestran la existencia de un tipo de docencia en el doctorado de carácter incidental, no coherente y principalmente basada en los intereses de investigación de los profesores. Aunque no disponemos de datos que permitan la comparación con estos resultados, la demanda de nuestros alumnos se orienta en el mismo sentido que recogen otras investigaciones: mayor formación en metodología, técnicas de investigación, diseño, organización y redacción de un proyecto de investigación, etc.

Por lo que respecta a los aspectos personales (motivación y habilidades), no parecen existir diferencias importantes entre disciplinas. Prácticamente todos los estudiantes tienen intención de realizar la tesis doctoral, no encuentran grandes dificultades o inconvenientes para seguir el tercer ciclo, perciben un nivel de exigencia moderado, parten de expedientes académicos con promedios similares¹⁰ y se dedican al doctorado más o menos en la misma medida. Quizá sorprende el número de horas de dedicación total al tercer ciclo, promedio de 15,77 horas semanales, frente a las 44,8 horas de trabajo personal y asistencia a clase dedicadas por los estudiantes estadounidenses (Gillingham, Seneca y Taussig, 1991). Nuestros estudiantes, por el contrario, trabajan más horas en un empleo remunerado (35,35 horas semanales por 16,3 horas de los alumnos americanos). A pesar de las limitaciones impuestas por los distintos procedimientos utilizados para la recogida de datos y las características y exigencias administrativas y académicas de ambos contextos, sería interesante profundizar en estos datos.

En otro orden de cosas, merece la pena destacar las distintas razones y objetivos que han motivado la realización del tercer ciclo en uno y otro colectivo de estudiantes de las áreas disciplinarias consideradas en el estudio. Parece que en humanidades y ciencias sociales se busca formación continua y mantener la vinculación con la universidad en mayor medida que en el área de ciencias. La satisfacción de estos objetivos no exige la elaboración de la tesis doctoral, lo que podría explicar el abandono del programa cuando el alumno se da cuenta de que se alarga el tiempo fijado al inicio para completar la tesis (Gillingham, Seneca y Taussig, 1991).

Aunque algunos de los datos expuestos puedan ser conocidos, o incluso intuitivamente previsibles, creemos conveniente que se lleven a cabo estudios sistemáticos sobre el tercer ciclo universitario en nuestro contexto educativo, empezando por los aspectos más descriptivos, que permitan trazar las líneas de lo que deberán ser análi-

10 Las notas de licenciatura no parecen estar relacionadas con el progreso en el doctorado (Pyke y Sheridan, 1993).

sis posteriores de los procesos de enseñanza, supervisión, orientación al estudiante, etc., sobre los que basar las decisiones administrativas y educativas.

Los resultados de este trabajo constituyen un primer análisis, parcial y no generalizable, acerca de las diferencias entre disciplinas académicas en relación con algunos factores influyentes en el éxito del tercer ciclo universitario. No nos resistimos, sin embargo, a plantear algunas consideraciones, recogidas en la literatura sobre el tema, que quizá podrían mejorar el rendimiento de los programas de doctorado, sobre todo en ciencias sociales y humanidades:

Vincular al menos algunas materias o cursos de doctorado con posibles temas de tesis (Hamilton, 1993).

Revisar aspectos como la extensión de las tesis, la cantidad de contribución original, el formato, o su posible «esponsorización» (Holdaway, 1994).

Permitir la coautoría de tesis (Hamilton, 1993).

Evaluar el rendimiento del profesorado encargado de supervisar tesis, en relación con el progreso realizado por el estudiante (Hamilton, 1993).

Establecer contratos de supervisión entre director y estudiante con el fin de clarificar compromisos y responsabilidades (Hockey, 1996 a).

Fomentar el trabajo cooperativo y la colaboración en materia de investigación en los departamentos universitarios (Anderson, 1996), especialmente en humanidades y ciencias sociales.

Proporcionar mayor formación metodológica e investigadora relacionada con los proyectos de investigación de los estudiantes (existe controversia sobre la conveniencia de hacerlo a través de los cursos de doctorado (Hockey, 1991).

Estimular la orientación al estudiante (y a los directores) y establecer programas de apoyo para la realización de la tesis (y para su supervisión) (Willcoxson, 1994).

Promover y facilitar la integración de los estudiantes lo más pronto posible en los proyectos de investigación de los departamentos.

Incentivar la participación y el compromiso de los profesores en el tercer ciclo.

Revisar y flexibilizar (según disciplinas) los objetivos, contenidos y exigencia de los cursos de doctorado y su adecuación a la formación investigadora y docente de los estudiantes.

Promover la redacción de un esquema del proyecto de investigación —por parte del supervisor— con los objetivos claramente detallados y las razones de la elección del tema, al inicio del programa (Hockey, 1991).

REFERENCIAS BIBLIOGRÁFICAS

- Abedi, J. y Benkin, E. (1987). The effects of students' academic, financial, and demographic variables on time to the doctorate, *Research in Higher Education*, 27, 3-14.
- Anderson, M.S. (1996). Collaboration, the doctoral experience, and the departmental environment, *The Review of Higher Education*, 19, 305-326.
- Baird, L.L. (1990). Disciplines and doctorates: The relationships between program characteristics and the duration of doctoral study, *Research in Higher Education*, 31, 369-385.

- Ballard, B. y Clanchy, J. (1993). Supervising students from overseas. En D.J. Cullen (Ed.), *Quality in PhD education* (pp. 61-73). Canberra: Australian National University.
- Becher, T., Henkel, M. y Kogan, M. (1994). *Graduate education in Britain*. Londres: Jessica Kingsley.
- Bowen, W.G. y Rudenstine, N.L. (1992). *In pursuit of the Ph.D.*, Princeton, N.J.: Princeton University Press.
- Bowen, W.G. y Sosa, J.A. (1989). *Prospects for Faculty in the Arts and Sciences*, Princeton, N.J.: Princeton University Press.
- Dolph, R.F. (1983). *Factors relating to success or failure in obtaining the doctorate*. Tesis doctoral no publicada. Universidad del Estado de Georgia.
- Donald, J.G., Saroyan, A. y Denison, D.B. (1995). Graduate student supervision policies and procedures: A case study of issues and factors affecting graduate study, *The Canadian Journal of Higher Education/ La Revue Canadienne d'Enseignement Supérieur*, 25, 71-92.
- Duggan, J.J. (1989). Time to the doctoral degree and success rates: The Berkeley experience. Berkeley: Universidad de California en Berkeley (documento no publicado).
- Filteau, C. (Ed.) (1992). *Graduate graduation rates and time-to-completion: Colloquium proceedings*. Toronto, Ontario, Canadá: Ontario Council on Graduate Studies.
- Fletcher, J. y Stren, R. (1992). Discussion of the factors influencing time-to-completion in graduate programmes: Student views. En C. Filteau (Ed.), *Graduate graduation rates and time-to-completion: Colloquium proceedings*. Toronto, Ontario, Canada: Ontario Council on Graduate Studies.
- Gillingham, L., Seneca, J.J. y Taussig, M.K. (1991). The determinants of progress to the doctoral degree, *Research in Higher Education*, 32, 449-468.
- Girves, J.E. y Wemmerus, V. (1988). Developing models of graduate student degree progress, *Journal of Higher Education*, 59, 163-189.
- Hamilton, R.G. (1993). On the way to the professoriate: The dissertation, *New Directions for Teaching and Learning*, 54, 47-56.
- Hill, T., Acker, S., y Black, E. (1994). Research students and their supervisors in education and psychology, en R.G. Burgess (Ed.) *Postgraduate education and training in the Social Sciences: processes and products*. Londres: Jessica Kingsley.
- Hockey, J. (1991). The Social Science PhD: a literature review, *Studies in Higher Education*, 16, 319-332.
- Hockey, J. (1994). Establishing boundaries: problems and solutions in managing the PhD supervisor's role, *Cambridge Journal of Education*, 24, 293-305.
- Hockey, J. (1996a). A contractual solution to problems in the supervision of PhD degrees in the UK, *Studies in Higher Education*, 21, 359-371.
- Hockey, J. (1996b). Motives and meaning amongst PhD supervisors in the social sciences, *British Journal of Sociology of Education*, 17, 489-506.
- Holdaway, E.A. (1994). Organization and administration of graduate studies in Canadian universities, *The Canadian Journal of Higher Education/ La Revue Canadienne d'Enseignement Supérieur*, 24, 1-29.

- Holdaway, E., Deblois, C. y Winchester, I. (1995). Supervision of graduate students, *The Canadian Journal of Higher Education/ La Revue Canadienne d'Enseignement Supérieur*, 25, 1-29.
- Moses, I. (1985). *Supervising postgraduates*. Sydney, Australia: Higher Education Research and Development Society of Australasia.
- Moses, I. (1992). Good supervisory practice. En I. Moses (Ed.), *Research training and supervision* (pp. 11-15). Actas de la conferencia organizada por la Australian Vice-Chancellor's Committee y la Australian National Board of Employment, Education and Training. Camberra: Australian Research Council.
- O'Brien, P.K. (1995). The reform of doctoral dissertations in humanities and social studies, *Higher Education Review*, 28, 3-19.
- OCDE (1987). *Post-graduate education in the 1980s*. París: Autor.
- Parry, O., Atkinson, P. y Delamont, S. (1994). Disciplinary identities and doctoral work, en R.G. Burgess (Ed.) *Postgraduate education and training in the Social Sciences: processes and products*. Londres: Jessica Kingsley.
- Pearson, R., Seccombe, I., Pyke, G. y Connor, H. (1993). Employer demand for doctoral social scientist?, *Studies in Higher Education*, 18, 95-104.
- Pike, S.W. y Sheridan, P.M. (1993). Logistic regression analysis of graduate student retention, *The Canadian Journal of Higher Education*, 23, 44-64.
- Salmon, P. (1992). *Achieving a PhD: ten students' experience*. Oakhill, Stoke-on-Trent, England: Trentham Books.
- Sheridan, P.M. y Pyke, S.W. (1994). Predictors of time to completion of graduate degrees, *The Canadian Journal of Higher Education*, 24, 68-88.
- Smart, J.C., Baird, L.L. y Bode, R. (1991, noviembre). Discipline differences in the learning demands of doctoral programs. Comunicación presentada en la *reunión anual de la Association for the Study of Higher Education*, Boston.
- Smart, J.C. y Hagedorn, L.S. (1994). Enhancing professional competencies in graduate education, *The Review of Higher Education*, 17, 241-257.
- Stirati, A. y Cesaratto, S. (1995). The Italian Ph.D ten years on: Educational, Scientific and occupational outcomes, *Higher Education*, 30, 37-61.
- Townsend, B.K. y Mason, S.O. (1990). Career paths of graduates of Higher Education doctoral programs, *The Review of Higher Education*, 14, 63-81.
- Tuckman, H.P., Coyle, S. y Bae, Y. (1990). *On time to the doctorate*, Washington, DC: National Academic Press.
- Valarino, E. (1992). *Todo menos investigación*. Caracas: Equinoccio (2ª ed., 1994).
- Valarino, E., Meneses, R., Yáber, G. y Pujol, L. (1996). Optimización de la gerencia de investigaciones en postgrado, *Revista Latinoamericana de Psicología*, 28, 63-82.
- Willcoxson, L. (1994). Postgraduate supervision practices: Strategies for development and change, *Higher Education Research and Development*, 13, 157-166.
- Wright, J. y Lodwick, R. (1989). The process of the PhD: a study of the first year of doctoral study, *Research Papers in Education*, 4, 22-56.
- Yeates, M. (1991). *Doctoral graduation rates in Ontario universities*. Toronto, Ontario, Canada: Council of Ontario Universities/Ontario Council on Graduate Studies.

CONCORDANCIA Y ERRORES (TIPO I Y II) EN EL DIAGNÓSTICO LOGOPÉDICO PRIMARIO¹

Antonio Fernández Cano
Mariano Machuca Aceituno
Universidad de Granada

RESUMEN

Este estudio muestra la alta concordancia entre el diagnóstico logopédico primario (test de detección) realizado por maestros/as con el diagnóstico emitido por expertos en base a la utilización de un instrumento de medida ad hoc (test diagnóstico).

Adicionalmente, se exponen los bajos errores (tipo I o II) cometidos al utilizar tal test de detección, se explora la validez del instrumento (métrica y diagnóstica) y se indican qué trastornos específicos (variables discriminantes) permiten detectar a sujetos con déficits logopédicos.

ABSTRACT

This study shows the high accordance between the primary logopedical diagnosis (detection test) made by a teacher and a technical diagnosis issued by experts using an ad hoc instrument of measurement (diagnosis test).

Additionally, it exposes the low rates of error (type I and II) derived from the use of such detection test. The validity of this instrument (metric and diagnostical) is explored. Finally it also points out which specific anomalies (discriminant variables) allow us to detect subjects with logopedical deficits.

1 Con la colaboración de Encarnación Guerrero Roldán, maestra-logopeda.

INTRODUCCIÓN

La normalidad logopédica es una cuestión harto controvertida y su diagnóstico podría ser menos fácil de lo que a *prima facie* pareciera. Se viene admitiendo, sin un fuerte fundamento conceptual, que algunos escolares presentan síntomas de trastornos en el habla, que esos síntomas son categorizables y que, en última instancia, el sujeto con un repertorio verbal correcto es distinguible del que no lo tiene. Así, los maestros y maestras de enseñanza infantil y primaria emiten diagnósticos primarios, basándose en la observación consuetudinaria, llevados de su intuición personal, sobre el estadio/capacidad logopédica de sus alumnos.

En el contexto andaluz, son los docentes quienes primero diagnostican de un modo global y genérico un posible trastorno logopédico en sus alumnos. Esos mismo alumnos serán posteriormente observados, diagnosticados y tratados, si procede, por un equipo logopédico-pedagógico. El diagnóstico primario (por primero, simple y proto-empírico) del maestro (primera aproximación al diagnóstico por anamnesis) se confronta con el diagnóstico *científico*, podríamos decir, del experto, produciéndose una situación similar a la toma de decisión en un contraste de hipótesis.

Tabla 1

TABLA DE CONTINGENCIA ENTRE DIAGNÓSTICOS (PRIMARIO *vs.* EXPERTOS)

	DIAGNÓSTICO PRIMARIO		
DIAGNÓSTICO	Normal	Normal	Anómalo
		Concordancia	Error Tipo I
EXPERTOS	Anómalo	Error Tipo II	Concordancia

Habría que esperar una alta concordancia entre sendos diagnósticos, aunque los errores a cometer producto de la discordancia, también son posibles.

El problema que estamos considerando no es nada nuevo, en Psiquiatría tiene una larga tradición y amplia consideración cuando se cotejan diagnósticos personales («me siento mal, doctor,») o sociales («X está necesitado de asistencia psiquiátrica») frente a diagnósticos profesionales o de expertos. En este sentido es antológico, el experimento de Rosenham (1973). Este autor y sus colaboradores simulaban tener trastornos psiquiátricos («oían voces»), solicitando asistencia médica. En la mayoría de los casos fueron diagnosticados como «insanos» y necesitados de tratamiento específico. Es evidente, en este caso, la manifestación de un alto error tipo I, al afirmar la existencia de diferencias o anomalías cuando en verdad no las hay, pues lo sujetos estaban palmaria-mente normales.

Dentro de las ciencias médicas, la concordancia en el diagnóstico es una cuestión capital dado que, en multitud de ocasiones, no existen pruebas diagnósticas que determinen con certeza absoluta la presencia o ausencia de enfermedad proporcionando un diagnóstico definitivo. Pero aunque suene paradójico, la validez de una metodología en el diagnóstico se establece en relación a un método de referencia, a partir de criterios básicos de validez: sensibilidad o capacidad para detectar a los individuos enfermos, evitando la presencia de falsos negativos, y especificidad, o capacidad para descartar a los exentos de la enfermedad, evitando la presencia de falsos positivos. Para una valoración de pruebas diagnóstico puede localizarse en Rodríguez Tuñas y Martínez de la Concha (1989). Y aunque esta cuestión no es el interés primordial del presente trabajo, sí será sometida a cierta consideración.

Nuestros médicos de atención primaria cometen en abundancia este tipo de error cuando remiten a sus pacientes «hipocondríacos» a especialistas para un segundo y más profesional diagnóstico; y es que en ciencias médicas, como ya es bien sabido, cometer un error tipo II (diagnosticar como sano a una persona no sana) es mucho más grave que el error tipo I. De aquí, la abundancia de este último.

La bibliografía sobre errores tipo I y II en investigación en ciencias sociales es extensa y prolija, véanse Darr (1987), Greenwald (1975) o Lykken (1968), como las realizaciones más consistentes. En Fernández Cano (1995, 79-82) se expone una amplia relación de trabajos que denotan la existencia del *sesgo pro estudios positivos*: los estudios deben poner de manifiesto resultados estadísticamente significativos (positivos por rechazar la hipótesis nula, H_0) como un prerrequisito de calidad científica.

Otra manifestación sutil de este sesgo en el diagnóstico sería la «profecía de autocumplimiento», expuesta por Rosenthal y Jacobson (1968), en el sentido de que la expectativa (primer diagnóstico) se ve corroborada con su cumplimiento (segundo diagnóstico).

De hecho, en este estudio estamos enfrentando dos métodos de escrutinio: un test de detección, realizado por el maestro (diagnóstico por anamnesis), menos seguro, menos articulado, menos costoso y más fácil de ejecutar, menos preciso y, sobre todo, que no constituye la base para un tratamiento posterior, frente a un test diagnóstico, administrado por un experto, más preciso, caro y difícil de ejecutar y que, ante todo, constituye la base para un tratamiento logopédico posterior.

II. MÉTODO

II.1. El problema

Este estudio tratará de dar respuesta a las tres cuestiones siguientes:

- ¿Qué grado de concordancia existe entre el diagnóstico logopédico primario realizado por maestros tutores y el diagnóstico profesional realizado por expertos?
- ¿Qué errores tipo I y II se comenten al realizar un diagnóstico primario (primera aproximación al diagnóstico por anamnesis)?

- ¿Qué variables permiten discriminar entre sujetos con déficits logopédicos diagnosticados primariamente?

Adicionalmente, un objetivo de este estudio, como paso intermedio para responder a las cuestiones que se plantean, será elaborar un instrumento de evaluación global del lenguaje espontáneo.

II.2. Diseño

Se trata de un estudio exploratorio descriptivo en el que se utilizarán criterios de validación y el análisis discriminante como diseño de investigación y técnica de análisis de datos, al par. La estructura esquemática del diseño sería:

$$\begin{array}{l} M: G_1 X_1 X_2 \dots X_n \longrightarrow Y_1 \\ M: G_2 X_1 X_2 \dots X_n \longrightarrow Y_2 \end{array}$$

Consta pues de dos grupos previamente emparejados (M: *Matched*) en base a una serie de variables intervinientes: edad, sexo y centro (los mismos valores para cada par de sujetos). Metodológicamente, el diseño sería una extensión ajustada del estudio de gemelos (*twin study*); véase la actualización que realiza Fischbein (1997) de este viejo método renacido de sus cenizas cual «ave fénix». La variable criterio (Y), diagnóstico primario del tutor/a, se dicotomiza en dos niveles: el nivel Y_1 , sujetos diagnosticados con anomalías y que configuran el grupo G_1 , y el nivel Y_2 , sujetos que no presentan déficits logopédicos a juicio de su tutor/a y que se encuadran en el grupo G_2 .

Las variables: $X_1 X_2 \dots X_n$ medidas en cada grupo serán las posibles variables discriminantes a dilucidar.

3. Variables

La variable criterio (Y) ya se ha definido como diagnóstico logopédico primario, mensurable a nivel dicotómico.

Las variables prodiscriminantes de interés serían una serie de parámetros relativos a los posible trastornos lingüísticos, diversificadas según:

A. *Nivel elocutivo o de codificación fonológica-sintáctica*, manifestable en:

1. *Parafasias fonémicas*: fallos en la producción de los fonemas o en el encadenamiento de los mismos para formar palabras. Ejemplo: «beno» (bueno), «rifo» (grifo).

2. *Parafasias fonéticas*: errores en la realización de un fonema por fallo en los rasgos del mismo (*traits*). No se produce sustitución del sonido sino anomalías en los tiempos de emisión del mismo. Ejemplo: todos los casos de transformación fonética.

3. *Parafasias morfélicas* (morfológica o sintagmática): se manifiesta mediante cambios en la estructura de la palabra con independencia del significado de la misma. Ejemplo: «clavel» (claro), «caelas» (escaleras).

4. *Neologismos*: se trata de sustitución profunda en una palabra llegando a crear un término nuevo, alejado del original y extraño al idioma. Ejemplo: «cala» (guitarra), «pavesa» (abeja).

5. *Agramatismos*: supresión de flexiones gramaticales en género, número y/o tiempo verbal. Ejemplo: «un reyes» (un rey), «ayer voy» (ayer fui).

6. *Tanteos verbales*: son facilitaciones que realiza el propio sujeto para obtener una denominación completa y, a veces, correcta del vocablo. Ejemplo: «ga..ca..ca..racol» (caracol), «agri..ba..ero..jar..dinero» (jardinero).

B. *Nivel semántico*, manifestable en:

7. *Anomias*: se trata de la dificultad o imposibilidad para encontrar la palabra adecuada en la práctica lingüística oral. Se manifiesta en la definición por el uso, el empleo de perífrasis y/o circunloquios. Ejemplo: «ua..ua» (lobo), «hace le gesto de encender una cerilla» (cerilla), «para echar agua» (botijo).

8. *Parafasias verbales paradigmáticas* (semánticas): consistentes en la sustitución de una palabra por otra que está relacionada por el significado (relación en la red semántica) y afecta a toda su estructura. Ejemplo: «jamón» (queso), «gusano» (caracol).

9. *Parafasias monémicas*: cuando la transformación afecta a un monema (mínima unidad portadora de significado). Ejemplo: «sacapuntas» (sacacorchos), «secadora» (secamanos).

10. *Perseveraciones*: se trata de repeticiones de contenidos semánticos producidos anteriormente. Se refleja en la iteración de conceptos/términos ya enunciados y que el sujeto es incapaz de inhibir. Así, ante distintas imágenes que se le presentan para que las nombre, el sujeto nombra la última con los mismos términos que la primera.

11. *Latencia*: tiempo que tarda el sujeto en procesar la información que se le ofrece para que sea emitida; se mide en segundos. Cuando al sujeto se le presenta un objeto puede ocurrir que se tome un compás de espera que puede acabar en la producción o no. Este compás de espera o latencias se interpreta como una manifestación inhibitoria de la actividad nerviosa superior.

12. *Fatigabilidad*: se trata de una manifestación inhibitoria patológica de la actividad nerviosa superior registrable mediante la aparición de bostezos, distractibilidad, agresividad, hiperactividad motora y/o logorrea. Se mide a nivel nominal-dicotómico [1: No presencia/ 2: Sí presencia].

II.4. Instrumentos

Para medir la variable criterio se ha utilizado el juicio del profesor al considerar si un alumno necesitaba ayuda para mejorar su expresión verbal. En paralelo, se le demandaba a ese mismo profesor que indicase otro alumno de la misma edad, sexo y centro o aula que, a su juicio también, no presentase trastornos en el habla.

Para medir las variables prodiscriminantes se ha utilizado un instrumento para la evaluación del lenguaje espontáneo-intencional (en adelante, IELEI) elaborado por Machuca, González y Fernández Cano (1997).

Este instrumento trata de detectar las anomalías de la expresión en aspectos parciales de la expresión-comprensión (variables prodiscriminantes), situando al sujeto competencialmente en su oportuno nivel de codificación lingüística.

No disponemos en español de instrumentos que evalúen ampliamente el desarrollo fonemático, semántico, morfosintáctico y de vocabulario. Aportes parciales para la evaluación de tal área problemática y tenidos presente en la génesis de este instrumento serían los trabajos de Bosch (1984), centrado en el desarrollo fonológico, Serra (1984), relativo a la articulación. Monfort y Suárez (1989) elaboraron un registro fonológico baremado con población española que registra las peculiaridades del habla del niño y la compara con la producción promedio del grupo de niños de su edad. Otra prueba que explora el lenguaje oral es la PLON (Prueba de Lenguaje Oral de Navarra) de Aguinaga et al. (1989), basada en los trabajos de Bloom y Lahey (1978), y centrada sobre forma, contenido y uso del lenguaje. Se trata de una prueba relativa a grupo.

Podíamos seguir enunciando instrumentos de carácter logopédico pero todos ellos tienen, a nuestro entender, profundas limitaciones en el sentido de que:

- La mayoría son producto de traducciones anglosajonas, con lo que están culturalmente sesgados ya que como es bien sabido el español y el inglés no sólo difieren en su sistema fonológico, si no también en su estructura sintáctica, semántica y pragmática.

- Son de referencia a norma-grupo, por lo que no permiten ni indican patrones de errores susceptibles para elaborar programas terapéuticos concretos.

- No suelen contemplar la información que se tiene sobre adquisición morfológica, de manera que se aplican sin distinción a todos los niños, de todas las edades, las mismas formas verbales.

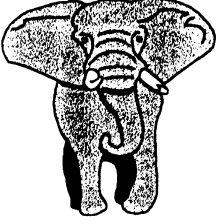


- Contemplan aspectos logopédicos parciales.

Por ello, nos pareció conveniente evaluar el estadio fonológico del lenguaje espontáneo (nivel intencional) y los procesos analítico-sintéticos inherentes al mismo a un nivel métrico. Mas, por otra parte, estimamos conveniente detectar ciertos parámetros de la comprensión que comprometen la extracción de lo significados, orientando el diagnóstico de sujetos con patología anártrica y/o afásica. En síntesis, el instrumento permite indicar el nivel de desempeño lingüístico a nivel codificativo, fonológico y semántico.

Descripción del instrumento IELEI

La prueba, tipo no verbal-verbal, trata de indagar diversos sistemas de conexiones entre el primer sistema (imágenes) y el segundo sistema de señales (lenguaje oral). En ella, el sujeto debe nombrar el objeto (imagen) que el evaluador le presenta.

La prueba consta de 150 imágenes-items, a modo de lotó, extraídas del Vocabulario Básico Curricular de Infantil y Primaria elaborado por Pérez González (1988), en la que están representados todos los elementos del código fonológico y sus combinaciones sintéticas en distintas posiciones de complejidad creciente. Para representar la prueba se sigue el criterio de desarrollo cronológico fonemático expuesto por Jakobson (1976). Véanse una imágenes-ejemplo:

		
elefante	martillo	gafas

Instrucciones para la administración

Se le informa al sujeto de la presentación de una serie de dibujos sencillos a los que tendrá que llamar por su nombre. Igualmente se le instruye a que debe anteponerle el determinante adecuado: un/una; el/la; uno/unas; los/las. Si alguna imagen no es reconocida debe manifestarlo para facilitársela bien fonológica o semánticamente.

En la hoja de respuestas se contemplan cinco apartados:

- Corrección o incorrección del vocablo emitido [SI/NO].
- Producción exacta del sujeto.
- Tiempo de latencia.
- Indicación si se oferta facilitación fonológica y/o semántica.
- Tipo de error/es si se ha/n cometido.

No existe límite de tiempo, aunque si el examinador observa fatiga por parte del sujeto se concluye la administración en una sesión posterior. El tiempo promedio ronda entre 25 y 30 minutos para alumnos de Primaria. En infantil, se completó en doble sesión de 15 minutos cada una. La administración fue individual, no existiendo límite de tiempo. La edad-suelo fue de cuatro años; la edad-techo, sin límite.

Validez y fiabilidad del IELEI

Dos aproximaciones a la validez del instrumento podríamos establecer: de constructo y de contenido. La validez de contenido estimamos que está garantizada pues queda contemplado representativamente a nuestro juicio el universo de trastornos lingüísticos admisibles (variables prodiscriminantes, anteriormente anunciadas).

Una aproximación a la validez de constructo del instrumento se ha realizado mediante análisis factorial a los componentes (trastornos/variables) contemplados. La solución más «bondadosa» e interpretable ha sido la ortogonal rotada (varimax) que manifiesta la existencia de cuatro factores específicos; ello, denota la complejidad del constructo «trastornos logopédicos» a falta de un factor general.

Tabla 2
SOLUCIÓN FACTORIAL ORTOGONAL Y ROTADA (VARIMAX) PARA COMPONENTES DEL CONSTRUCTO TRASTORNOS LOGOPÉDICOS

VARIABLE	I	II	III	IV	h ²
1. Latencia	0.89	-	-	-	0.80
2. Anomia	0.88	-	-	-	0.80
3. Neologismos	0.66	0.54	-	0.32	0.83
4. Morfémicas	0.65	0.32	-	-	0.57
5. Agramatismos	0.59	0.51	-	0.27	0.68
6. Perseveración	-	0.73	0.43	-	0.76
7. Fonéticas	-	0.70	-	-0.26	0.61
8. Parafasias V.	0.28	-	0.75	-	0.66
9. Fonémicas	0.29	-	-0.65	-	0.57
10. Tanteos	—	-	-	0.92	0.86
Varianza	2.95	1.71	1.31	1.20	

Cargas < |0.25| → 0

Una interpretación plausible de los factores obtenidos sería:

Factor 1: Cuadro afásico mixto.

Factor 2: Si excluimos la variable perseveración obtendríamos un cuadro anártrico puro y potente.

Factor 3: Cuadro con potencia afásica pura.

Factor 4: Cuadro anártrico de escasa potencia.

Obsérvese la presencia de dos trastornos genéricos bien delimitados: afásicos y anártricos. Aventuramos que una solución factorial de segundo orden aportaría evidencia más contundente a esta conjetura.

Se ha calculado una fiabilidad *interevaluadores-codificadores* del siguiente modo: Dos evaluadores independientes observan, por una parte, la presencia o ausencia de errores, y por otra, toman nota fiel de la respuesta o ausencia de la misma, que el sujeto va produciendo, relativa a todos y cada uno de los ítems presentados. Los casos discrepantes (un mismo sujeto es calificado de manera distinta ambos calificadores) son resueltos conjuntamente, volviendo de nuevo a administrar al sujeto el ítem discrepante y tomando un acuerdo final.

En una primera administración el acuerdo entre evaluadores logró un porcentaje de concordancia entorno al 86% de los casos. Los casos en que se evidenciaron discrepancia fueron 16 (14 %).

Posteriormente se calculó un índice de fiabilidad por consistencia interna de unidades, utilizando la \hat{U} de Carmines (una extensión del \hat{O} de Cronbach). El valor obte-

nido de 0,78 expresa una fiabilidad adecuada y más para un instrumento como el IELEI, inicial, exploratorio y poco depurado.

II.5. Muestra

La muestra utilizada en este estudio ha estado formada por 57 pares de sujetos de entre cuatro y doce años, de escolarización normal (en aulas ordinarias), pertenecientes a ocho centros de Jaén y provincia. Cada emparejamiento fue realizado por los propios tutores en base al mismo sexo, edad y pertenencia a misma aula o centro.

La mecánica de la selección de la muestra venía condicionada por la iniciativa de los tutores que derivaban a los alumnos para que fuesen estudiados por los Equipos de Orientación Educativa (expertos), por creer que presentaban, a consecuencia de déficits lingüísticos, dificultades en el aprendizaje. Posteriormente, se le demandó a cada profesor que presentase un sujeto «gemelo», pero que no presentase tales déficits.

La variable «sexo» está representada mostrando valores de frecuencia en torno a 68 % de varones y 32 % de mujeres. Este dato constituye de por sí un hallazgo, ya que el porcentaje de alumnos varones diagnosticados primariamente por el tutor como necesitados de atención logopédica es sensiblemente (con significación substantiva, a nuestro juicio) más alto que el de mujeres, en una proporción de 2:1.

En la variable «fetría» cabe reseñar que el 43 % de la muestra tiene al menos un hermano mayor. Este hecho es digno de destacar ya que clarifica especulaciones y puntos de vista diferentes acerca de que la alteración de la elocución suele ser afectar a niños unigénitos o que ocupan el primer lugar fetrico.

La composición de la muestra viene dada en la tabla III, teniendo en cuenta que cada valor representa un par de sujetos.

Tabla 3

DESCRIPCIÓN DE LA MUESTRA UTILIZADA EN EL PRESENTE ESTUDIO

EDAD	-os.	-as.	f_i	%
4	9	2	11	19%
5	12	6	18	32%
6	3	5	9	16%
7	6	2	8	14%
8	3	1	4	7%
9-12	5	2	7	12%
f_i	39	18	57	
%	68%	32%		100%

III. RESULTADOS

Para analizar los datos disponibles se ha utilizado la técnica multivariada del análisis discriminante (AD). Esta técnica ha estado estrechamente asociada al problema de la clasificación dado que un espacio discriminante podría ser óptimo para emitir una decisión de clasificación en ciertas situaciones (ejem. Roulon et al. 1967, pp. 299-319, 339-341). El AD sirve básicamente para aportar intuiciones sobre cómo grupos preformados (en este caso, según el criterio primario del tutor) difieren en un espacio métrico multivariado. El uso del AD en procesos de clasificación es antiguo, Bernard (1935) realizó su primera aplicación y desarrollo como técnica analítica tratando de detectar cuales eran las variaciones en cuatro series distintas de momias egipcias. Rao y Slater (1949) lo utilizaron para diferenciar entre grupos de neuróticos. Desde entonces, la técnica multivariada del AD ha tenido una fuerte implantación y desarrollo en ciencias sociales y conductuales.

III.1. Variables discriminantes

Como función informativa previa aportamos los resultados a nivel de grupo parcial (sujetos diagnosticados como anómalos frente a normales) en las variables que se consideraron prodiscriminantes a nivel de medias aritméticas.

Tabla 4

VALORES PROMEDIO DE LA PRESENCIA DE TRASTORNOS LOGOPÉDICOS
DETECTADOS POR EL IELEI

VARIABLE/ TRASTORNO	G. anómalos (A) (n = 57)	G. normales (N) (n = 57)
Fonémicas	19.2	0.0
Fonéticas	3.8	0.0
Morfémicas	2.7	0.0
Neologismos	9.3	0.0
Tanteos	2.8	0.1
Agramatismos	7.8	0.1
Paradigmáticas	9.4	17.1
Anomias	11.4	7.9
Latencias	8.1	1.5
Perseveraciones	1.1	0.0
Monémicas	0.0	0.0
Fatigabilidad	1.2	1.0

Se denotan diferencias entre grupos en la mayoría de las variables, y en la dirección plausible: más errores en el grupo A que en el N. Llama poderosamente la atención la variable «parafasias verbales *paradigmáticas*», en el sentido de que no se verifica la dirección que inicialmente intuíamos.

Aplicando el AD a los datos anteriores y tras un proceso en cuatro pasos, las variables que permiten discriminar la pertenencia de un caso/sujeto a un grupo u a otro, son las dadas en la siguiente tabla:

Tabla 5

VARIABLES DISCRIMINANTES DE LA DICOTOMIA «ANOMALIA-NORMALIDAD» LOGOPÉDICA SEGÚN DIAGNÓSTICO PRIMARIO

Paso	Variable	F	L de Wilks	F	g.l.
1	Fonémicas	96.3	0.537	96.3	1;112
2	Paradigmáticas	13.8	0.479	60.5	2;111
3	Fatigabilidad	15.3	0.419	50.7	3;110
4	Agramatismos	6.0	0.397	41.3	4;109

III.2. Concordancia entre diagnósticos

La siguiente tabla de contingencia (tabla 6), en la que aparecen los sujetos clasificados en cada grupo y el porcentaje correspondiente, nos atestigua de una fuerte concordancia entre ambos tipo de diagnóstico: el primario del tutor y el «científico» de los expertos.

Tabla 6

TABLA DE CONTINGENCIA CON CASOS (Y PORCENTAJES) DE ACUERDO Y ERRORES TIPO I Y II

DIAGNÓSTICO EXPERTOS	DIAGNÓSTICO PRIMARIO			
		Normal	Anómalo	
Normal		56 (.49) VP	8 (0.07) FP	64
	Anómalo	1 (.01) FN	49 (.43) VN	50
		57 (.50)	57 (.50)	114

En el 92 % (.49 + .43) de los casos hay concordancia entre ambos diagnósticos; incluso, adoptando un procedimiento clasificatorio más conservador cual es el del corte de navaja (*jackknife classification*), sólo aumentaría la clasificación en dos casos discrepantes, pasando de 8 a 10 los sujetos clasificados primariamente como anómalos cuando según los especialistas son normales. Tan alta concordancia nos hace tener bastante confianza en el sentido detector del tutor. De ahí la conveniencia de formular el «diagnóstico por anamnesis» en un contraste de hipótesis que el técnico del diagnóstico tendrá que confirmar o rechazar (véase Machuca, 1996).

III.3. Parámetros en la concordancia diagnóstica

Utilizando conceptualizaciones de la epidemiología, disciplina médica preocupada por cuestiones diagnósticas (i.e. Ahlbom, 1987; Janicek y Cloeroux, 1987; Rothman, 1987), podríamos obtener ciertos parámetros afines de especial relevancia. A partir de la tabla anterior y sabiendo que los casos anteriores configuran la tipología: VP (verdaderos positivos), VN (verdaderos negativos), FP (falsos positivos), FN (falsos negativos), calcularemos:

-Sensibilidad [S] del test de detección (diagnóstico primario): $S = VP / (VP + FN) = 0.89$.

-Especificidad [E] de dicho test: $E = VN / (FP + VN) = 0.87$.

El test de detección del tutor es más sensible que específico, o sea el maestro/a detecta a la mayoría de niños/as con anomalías logopédicas, pero también suele presentar como anómalos a sujetos que no lo son.

El valor global del test de detección [VG] podría calcularse: $VG = (VP + VN) / (VP + VN + FN + FP) = 0.91$, un índice bastante alto que denota la valía de tal procedimiento. Incluso, cuando se descuenta el posible acuerdo debido al efecto del azar, utilizando la K de Cohen, se obtiene un valor de 0.84, indicador muy aceptable como expresión de concordancia.

El valor predictivo del diagnóstico primario podría determinarse tanto para el test positivo (VP^+) como para el test negativo (VP^-). Así, $VP^+ = VP / (VP + FP) = 0.86$ y $VP^- = VN / (VN + FN) = 0.98$. Al ser estos dos parámetros bastante similares, podemos afirmar que la prevalencia de la «enfermedad» (trastornos logopédicos) es media a nivel poblacional ya que es bien sabido que cuanto menor sea la prevalencia de las anomalías menor será VP^+ y mayor VP^- , y viceversa en el caso de una alta prevalencia.

III.4. Tipo de errores

A partir de los datos disponibles se constata que el error tipo I (sujetos diagnosticados primariamente como anómalos pero que tras un riguroso examen son calificados normales) excede al error tipo II (sujetos diagnosticados como normales pero que necesitarían tratamiento logopédico según se constata en el examen posterior de los especialistas).

Se corrobora, una vez más, el fenómeno que denota una mayor abundancia del error tipo I (falsas diferencias en 8 casos) frente al de tipo II (falsas igualdades en 1 caso). O sea, casi un 1 % frente al 7 %. No obstante, en este contexto, no deja de ser pre-

ocupante ese caso único de error tipo II: un sujeto con apariencia de normalidad pero necesitado de tratamiento específico.

IV. CONCLUSIONES

Los hallazgos básicos que ofrece este estudio podríamos resumirlos como sigue:

— Existe una serie de variables que permiten discriminar la existencia de anomalías logopédicas en escolares de 4 a 12 años, en base a un instrumento de medida con cierta validez y fiabilidad. Tales variables son las relativas a la presencia trastornos fonémicos, paradigmáticos y agramatismos, así como la variable fatigabilidad.

— El sentido de «detección» de anomalías logopédicas en escolares que posee el profesor tutor concuerda con un diagnóstico más ajustado que emite el especialista con el concurso de un instrumento validado.

— Existen mínimas discordancias entre el diagnóstico primario del profesor y el acometido por especialistas apoyados en un instrumento de medida. Tales discrepancias expresables como errores expone una mayor tasa de error tipo I (falsas diferencias) que error tipo II (falsas igualdades) en una razón de 7:1.

V. RECOMENDACIONES

Apelar al manido «se necesita más investigación al respecto», aunque ésta sea de réplica, es el único modo de obtener una evidencia más contundente, que no definitiva, sobre las cuestiones que aquí se indagan.

Mejoras en los posibles estudios de réplica irían en conseguir grupos criteriosales lo más «gemelizados» posible que incorporen, aún a costa de más mortalidad, mayor número de variables de emparejamiento.

Igualmente, parece conveniente indagar en los mecanismos íntimos que el tutor utiliza como test de detección para diferenciar sujetos con o sin anomalías logopédicas. La dicotomización de «caja negra», que aquí hemos utilizado como variable criterial, merecería ser explorada en profundidad.

BIBLIOGRAFÍA

- Aguinaga, G.; Armentia, M.L.; Fraile, A.; Olangua, P. y Uriz, N. (1989). *Prueba de lenguaje oral Navarra (PLON)*. Pamplona: Fondo de Publicaciones del Gobierno de Navarra (Departamento de Presidencia-Interior).
- Ahlbom, A. y Norell, S. (1987). *Fundamentos de epidemiología*. Madrid: Siglo XXI.
- Barnard, M.M. (1935). The secular variations of skull characters in four series of egyptian skulls. *Annals of Eugenics*, 6, 352-371.
- Bloom, L. y Lahey, M. (1978). *Language development and language disorders*. Nueva York: John Wiley and Sons.
- Bosch, L. (1984). El desarrollo fonológico infantil: Una prueba para su evaluación. En *Estudios de psicología infantil*, M. Siguan (Ed.). Madrid: Pirámide.

- Darr, R. (1987). Another look at Meehl, Lakatos, and the scientific practices of psychologist. *American Psychologist*, 42, 145-151.
- Fernández Cano, A. (1995). *Métodos para evaluar la investigación en Psicopedagogía*. Madrid: Síntesis.
- Fischbein, S. (1997). Twin studies. En *Educational Research, Methodology, and Measurement: An International Handbook*, (2ª ed.), J.P. Keeves (Ed.), pp. 162-167. Oxford: Pergamón.
- Jakobson, R. (1976). *Lenguaje infantil y afasia*. Barcelona: Alianza.
- Greenwald, A. (1975). Consequences of prejudice against the null hypothesis. *Psychological Bulletin*, 81, 1-20.
- Jenicek, M. y Cleroux, R. (1987). *Epidemiología. Principios, técnicas y aplicaciones*. Barcelona: Salvat.
- Lykken, D. (1968). Statistical significance in psychological research. *Psychological Bulletin*, 70, 151-159. También en *The significance test controversy*, (1970) D. Morrison y R. Henkel (Eds.), pp. 267-269. Chicago: Aldine.
- Machuca, M. (1996). Hacia un enfoque comprensivo del diagnóstico en educación. *Jaenseñanza*, 10, 95-100.
- Machuca, M. y Fernández Cano, A. (1997). Elaboración de un instrumento para la evaluación del lenguaje espontáneo-intencional. (Documento interno). Granada: Facultad de Educación.
- Monfort, M. y Juárez, A. (1989). *Registro fonológico inducido*. Madrid: CEPE.
- Pérez González, J. (1988). *Didáctica de la Ortografía*. Valladolid: Miñón.
- Rao, C.R. y Slater, P. (1949). Multivariate analysis applied to differences between neurotic groups. *British Journal of Psychology, Statistical Section*, 2(1), 17-29.
- Rodríguez Tuñas, B. y Martínez de la Concha, D. (1989). Valoración de pruebas diagnósticas. En *Curso de Metodología de Investigación en Enfermería*, pp. 227-241. Granada: ICE-Universidad de Granada.
- Rosenham, D.L. (1973). On being sane in insane places. *Science*, 179, 250-259.
- Rosenthal, R. y Jacobson, L. (1968). *Pygmalion in the classroom*. Nueva York: Holt, Rinehart & Winston.
- Rothman, K.J. (1987). *Epidemiología moderna*. Madrid: Ed. Díaz de los Santos.
- Roulon, P.J.; Tiedeman, D.U.; Tatsuoaka, M.M. y Langmuir, C.R. (1967). *Multivariate Statistics for Personnel Classification*. Nueva York: Wiley.
- Serra, M. (1984). Normas estadísticas para la población de 3 a 7 años del área metropolitana de Barcelona. *Logopedia y Fonoaudiología*, 3-4, 232-235.

PROPUESTA METODOLÓGICA PARA LA EVALUACIÓN DEL GASTO EDUCATIVO DE LOS HOGARES EN EDUCACIÓN REGLADA NO UNIVERSITARIA. UNA APLICACIÓN A LA COMUNIDAD AUTÓNOMA DE LA REGIÓN DE MURCIA

Juan Gómez García, Matilde Lafuente Lechuga, Antonio Losa Carmona

Facultad de Ciencias Económicas y Empresariales
Universidad de Murcia

RESUMEN

En este trabajo se abordan diversas cuestiones relativas al gasto educativo de los hogares de la Región de Murcia. Por un lado, el estudio de las características de los hogares con miembros realizando estudios de primer y segundo grado. Por otro, una estimación del gasto educativo de los hogares de la Región de Murcia en el curso 1995/96, diferenciando gastos totales y por alumno, por niveles de enseñanza, y siempre que ha sido posible, según la titularidad del centro. Por último, haciendo uso de técnicas multivariantes, se determinan los factores fundamentales de que depende la variable gasto en enseñanza.

***Palabras clave:** gastos educativos, estudios primarios, estudios secundarios, economía de la educación, centros públicos, centros privados, técnicas multivariantes.*

ABSTRACT

This paper deals with several issues relative to the Murcia's households educational expenditure. On one hand, it is studied the characteristics of households with members undergoing primary and secondary studies. On the other, an estimation of the Murcia's households educational expenditure in the 95/96 course, taking into account total expenditures and per student,

in each educational level and, if possible, according to the ownership of the center. Finally, fundamental factors on which the educational expenditure variable depends are determined by means of multivariate techniques.

Key words: *educational expenditure, primary studies, secondary studies, economics of education, public education, private education, multivariate techniques.*

I. INTRODUCCIÓN

Pese a la naturaleza esencial de la educación, por su incidencia en múltiples esferas del progreso humano, tanto en el ámbito de las relaciones sociales y económicas, como en el más íntimo del desarrollo de la personalidad, el esfuerzo destinado en nuestro país a la elaboración de un sistema contable adecuado sobre consecuencias económicas de las actividades educativas ha sido nulo hasta fechas bien recientes. Esta carencia de información ya había sido denunciada hace ahora dos años por Palafox, Mora y Pérez (1995), al constatar la falta de explicaciones convincentes a cuestiones básicas relativas al qué, al cómo y al para quién se produce cuando se formulan en el ámbito propio de la educación.

En el empeño de dar respuesta pertinente a esta necesidad informativa, y cuando el presente trabajo había empezado a tomar forma para su presentación en la XXIII Reunión de Estudios Regionales, celebrada en Valencia en noviembre de 1997, aparece un estudio realizado por un equipo de investigadores de la Universidad de Valencia [Uriel, E. et al. (1997)], cuyo propósito fundamental, tal y como reza el título de su trabajo, es la elaboración de las cuentas de la educación en España y sus Comunidades Autónomas, en el período 1980 a 1992. Pero dejemos que sean los propios autores quienes determinen el alcance exacto de su aportación: «Las CE [cuentas de la educación], mediante el desarrollo de categorías adicionales y la información estadística correspondiente a nuevos conceptos que sobrepasan el núcleo del sistema central de las cuentas nacionales,... [nos permiten] expandir la capacidad analítica de las cuentas nacionales, estableciendo las conexiones necesarias entre diversos aspectos del fenómeno de la educación»¹.

El presente trabajo se inscribe en la línea de investigación iniciada por ese equipo de trabajo; si bien, se constatan algunas diferencias dignas de mención.

Por un lado, nuestro estudio tiene un alcance mucho más limitado que el de Uriel, por diversas razones. En razón del ámbito espacial, se confina exclusivamente a la Comunidad de Murcia, quedando, por tanto, fuera de su alcance el análisis económico de las actividades educativas desarrolladas en otros territorios autonómicos, incluso las ejercidas en el territorio nacional como un todo; en razón del sujeto agente analizado, se circunscribe al sector de hogares, identificado en el cuadro central de la Contabilidad Nacional en las unidades institucionales homónimas; y en fin, porque nuestras estimaciones se ciñen a los niveles reglados de la educación no universitaria: preescolar, EGB, Formación Profesional, BUP y COU.

1 Uriel, E.; Moltó, M.L.; Pérez, F.; Aldás, J.; y Cucarella, V. (1997, p. 24).

Por otro lado, como es lógico, nuestro estudio ofrece un grado de detalle superior al ofrecido en el trabajo colectivo de referencia, puesto que el gasto educativo se analiza bajo un número superior de categorías. En efecto, la caracterización de los hogares y del sustentador principal nos permite sobrepasar los 18 tipos de hogares que aparecen en el estudio de Uriel et al., ya que, además, de ofrecer un detalle superior en las categorías que son comunes en los dos estudios: las referidas al tamaño del municipio, o hábitat, nivel de ingresos del hogar, o estudios del sustentador principal; también se consideran otras categorías diferentes como el número de miembros y la principal fuente de ingresos del hogar, así como el sexo, edad, relación con la actividad, condición socioeconómica y categoría socioprofesional del sustentador principal del hogar, que nos ofrecen un perfil más preciso de los beneficiarios del servicio educativo.

Así pues, el objeto fundamental de este trabajo es la estimación del gasto en enseñanza primaria y secundaria realizado por los hogares murcianos en el curso 1994/95, y se estructura del siguiente modo: El epígrafe siguiente se destina al estudio de los aspectos metodológicos, ofreciendo asimismo un detalle de la naturaleza de las fuentes estadísticas utilizadas, especialmente nuestra atención se va a centrar en la Encuesta de Presupuestos Familiares de 1990/91. En el apartado 3 se analizan las características básicas de los usuarios de la educación, para ello se hace un estudio descriptivo bastante detallado de los hogares con miembros estudiando en estos niveles educativos. En el siguiente epígrafe se ofrece una estimación del gasto educativo de los hogares de la Región de Murcia en el curso 1995/96, diferenciando gastos totales y por alumno, por niveles de enseñanza, y siempre que ha sido posible, según la titularidad del centro. En el epígrafe 5 aparecen los resultados obtenidos de la aplicación de técnicas multivariantes que nos han servido para determinar los factores fundamentales de que depende la variable gasto en enseñanza. Por último, el apartado 6 recoge las conclusiones fundamentales del trabajo.

2. METODOLOGÍA

Los niveles educativos que vamos a tener en cuenta son los siguientes: Preescolar, ciclos inicial y medio de EGB, ciclo superior de EGB, BUP, COU, Selectividad y Pruebas de Acceso para mayores de 25 años, FP de 1º grado y FP de 2º grado.

Para abordar estos propósitos, la fuente estadística que proporciona una información más exhaustiva y fiable es la Encuesta de Presupuestos Familiares que elabora cada diez años, aproximadamente, el Instituto Nacional de Estadística, y cuyo objetivo fundamental es la elaboración de un nuevo Índice de Precios al Consumo. La última Encuesta Básica de Presupuestos Familiares tiene como período de referencia desde abril de 1990 a marzo de 1991. Otro tipo de investigación, que desarrolla también el mismo organismo, tiene un carácter uniforme en el tiempo, la Encuesta Continua de Presupuestos Familiares, y persigue conocer la evolución trimestral del gasto de los hogares y su distribución por rúbricas de gasto.

La situación ideal hubiese sido elaborar un estudio con base en la Encuesta de Presupuestos Familiares 90/91, y complementar esta información con la que se

deduce de la Encuesta Continua de Presupuestos Familiares, pero al no presentar esta última encuesta información por Comunidades Autónomas —esto parece que va a cambiar a lo largo del año 1997—, sólo podremos hacer uso de la última encuesta estructural.

El Instituto Nacional de Estadística nos ha suministrado en soporte magnético los ficheros deducidos de la Encuesta de Presupuestos Familiares 90/91 correspondientes a los hogares residentes en la Región de Murcia. De los cinco ficheros recibidos (datos generales del hogar, fichero de gastos, datos de los miembros del hogar, datos de los bienes de equipamiento y datos de las viviendas secundarias), hemos utilizado los tres primeros.

El fichero de datos generales del hogar, contiene, además de información sobre la composición del hogar, datos del sustentador principal, así como cifras de ingresos y gastos totales. Incluye un total de 526 registros, que se corresponden con el número de hogares residentes en la Región de Murcia de los que se consiguió información. En este archivo sólo aparece el gasto total del hogar, y no se puede deducir de aquí la naturaleza del mismo. Para ello, tenemos que utilizar el fichero tipo 2, que ofrece información detallada del gasto. Aquí aparecen clasificados los gastos según la clasificación denominada PROCOME, nomenclatura que consta de 7 dígitos, que coincide en sus cuatros primeros niveles con la Clasificación del Consumo final de las familias, según la nomenclatura armonizada en el seno de la Oficina de Estadística de las Comunidades Europeas.

Del fichero de gastos nos hemos limitado, como es obvio, a los registros que hacen mención a algún tipo de gasto educativo, o anexo al mismo, en los niveles considerados. Tal y como está estructurada la nomenclatura PROCOME, no todos los gastos se han podido clasificar atendiendo a la titularidad del centro educativo por centros privados o públicos; sólo los gastos regulares y de inscripción se han podido desglosar a este nivel de detalle; el resto de gastos se consideran comunes a los centros para cada nivel de enseñanza.

En resumidas cuentas, los gastos para los distintos niveles educativos, los hemos detallado atendiendo a la siguiente clasificación:

- *Gastos regulares y de inscripción:* como ya hemos hecho mención, este tipo de gastos se ha podido desglosar para centros privados y públicos.
- *Gastos comunes:* Se trata de los gastos que no se pueden detallar según la titularidad del centro. Entre ellos se encuentran los libros de texto, el material escolar, el seguro escolar, los gastos en transporte, en asociaciones de padres, en comedor escolar, en otras actividades, etc.

La necesidad de utilizar el fichero tipo 3, que contiene información sobre los miembros del hogar, se planteó por la imposibilidad de desglosar el gasto de determinadas rúbricas por niveles educativos; en concreto, para los gastos en material escolar, seguro escolar, transporte, comedor escolar y asociaciones de padres, entre otros.

El resultado final es un fichero de familias en el que cada hogar tiene asignado un vector de gastos y un vector de miembros. Sobre este fichero se ha hecho un estudio

descriptivo detallado que compara la distribución de hogares con algún miembro-alumno con la del conjunto de todos los hogares de la Región de Murcia. También se ha extraído, a partir de este archivo, la distribución de los alumnos por niveles educativos, atendiendo a ciertas características del hogar y del sustentador principal, y según estén matriculados en centros públicos o privados.

Una vez obtenidos los gastos de los dos tipos considerados para cada uno de los niveles educativos antes definidos, se han utilizado los datos que sobre alumnos matriculados publica el Centro Regional de Estadística en el Anuario Estadístico de la Región de Murcia. Se ha extraído la información relativa al curso 90/91, obteniéndose de esta forma los correspondientes gastos por alumno.

Para estimar el gasto educativo en el curso 95-96, hemos hecho la hipótesis de que durante los cinco años que median entre el año de la Encuesta y este curso, se ha mantenido la estructura de gasto de los hogares, de forma que para obtener una estimación del gasto por alumno en el curso 95/96, a pesetas de ese año, se ha deflactado el dato correspondiente al curso 90/91 haciendo uso del Índice de Precios al Consumo de Cultura, Enseñanza y Esparcimiento con base 1992, desde abril de 1990 hasta julio de 1996.

3. ANÁLISIS DESCRIPTIVO DE LOS HOGARES CON MIEMBROS CURSANDO ESTUDIOS DE PRIMER Y SEGUNDO GRADO

Haciendo uso de los factores de elevación poblacional, la muestra seleccionada para la Región de Murcia representa a un total cercano a 285.000 unidades. Según se deduce de los resultados de la encuesta, el 40% de estos hogares tienen algún miembro estudiando en los niveles considerados.

Nuestra atención se ha dirigido, en primer lugar, al estudio de tres características del hogar: tamaño del municipio de residencia, número de miembros y principal fuente de ingresos (tabla 1).

Más de la mitad de los hogares con gastos en enseñanza de primer y segundo grado tienen su residencia en Municipios mayores de 100.000 habitantes, que en el caso de nuestra Comunidad Autónoma son dos, Cartagena y Murcia. El otro grupo de Municipios que presenta mayor representación es el de aquéllos cuyo tamaño está comprendido entre 10.000 y 50.000 habitantes, que supone el 35%, mientras que los hogares residentes en poblaciones menores de 10.000 habitantes tienen una participación inferior. Si contrastamos esta distribución, con la del conjunto de los hogares, observamos unas estructuras bastantes similares, exceptuando, precisamente, los municipios de menor tamaño que presentan mayores diferencias.

El número de personas que componen el hogar debe ser, lógicamente, determinante del número de miembros cursando estudios del primer y segundo grado. El 75% de los hogares que cumplen esta condición tienen entre 4 y 6 miembros, en contraste claro con la participación de esta categoría en el total de los hogares que es del 48%.

Si clasificamos la población en estudio de acuerdo con la principal fuente de ingresos, hemos encontrado que el 72% de los hogares cumplen la condición de que sus

Tabla 1
DISTRIBUCIÓN POR ALGUNAS CARACTERÍSTICAS DEL HOGAR

	Total hogares (%)	Hogares con miembros estudiando (%)	% Hogares estudio / total hogares
TAMAÑO DEL MUNICIPIO DE RESIDENCIA			
Hasta 10.000 habitantes	9,3	6,1	26,5
De 10.000 a 50.000 habitantes	37,2	35,3	38,1
De 50.000 a 100.000 habitantes	6,2	6,7	43,5
Más de 100.000 habitantes	47,3	51,8	44,1
NÚMERO DE MIEMBROS DEL HOGAR			
Menos de 4 miembros	47,8	16,7	14,1
De 4 a 6 miembros	47,6	74,6	63,0
Más de 6 miembros	4,6	8,7	76,0
PRINCIPAL FUENTE DE INGRESOS DEL HOGAR			
Por trabajo por cuenta ajena	55,6	71,8	52,0
Por trabajo por cuenta propia	11,7	17,8	60,9
Por pensión y otros subsidios	31,8	8,7	11,1
Otros	0,9	1,6	75,6
TOTAL	100,0	100,0	40,2

ingresos provienen principalmente del trabajo por cuenta ajena, mientras que esta cuota es del 56% para el total de la población. La categoría definida por trabajo por cuenta propia, ya a mucha distancia, presenta participaciones del 18 y 12% respectivamente. Tendencia distinta presentan los hogares cuya principal fuente de ingresos es una pensión de jubilación o incapacidad u otros subsidios, pues de suponer más del 30% del total de la población, esta cuota se reduce a tan sólo el 9% de los hogares con miembros cursando estudios en los niveles considerados.

En relación con el sustentador principal del hogar, las características que se han estudiado han sido las siguientes: sexo, edad, relación con la actividad económica, nivel de estudios, condición socioeconómica y categoría socioprofesional.

Como era de esperar, el sexo del sustentador principal es mayoritariamente varón, llegando a suponer el 94% de los hogares con algún miembro realizando estudios de primer o segundo grado. Atendiendo a la edad del sustentador principal, la categoría que agrupa a los de 30 a 44 años es la que presenta una mayor participación, suponiendo casi la mitad de los hogares, seguida muy de cerca por la formada por la de edades comprendidas entre 45 y 64 años. Entre estos dos grupos se encuentran casi la totalidad de los hogares con gastos en enseñanza. Para el total de las unidades, el porcentaje de hogares con sustentador principal ocupado, es del 59%; pues bien, si nos

Tabla 2

DISTRIBUCIÓN POR ALGUNAS CARACTERÍSTICAS DEL SUSTENTADOR PRINCIPAL

	Total hogares (%)	Hogares con miembros estudiando (%)	% Hogares estudio / total hogares
SEXO DEL SUSTENTADOR PRINCIPAL			
Varones	84,2	94,3	45,0
Mujeres	15,8	5,7	14,5
EDAD DEL SUSTENTADOR PRINCIPAL			
De 18 a 29 años	8,0	5,5	27,5
De 30 a 44 años	30,2	47,5	63,2
De 45 a 64 años	40,8	44,9	44,2
De 65 y más años	21,0	2,2	4,1
RELACIÓN CON LA ACTIVIDAD DEL SUSTENTADOR PRINCIPAL			
Ocupados	59,1	84,0	57,2
Parados	5,1	5,1	40,3
Pensionistas	30,7	10,2	13,3
Otros inactivos	5,1	0,7	5,7
NIVEL DE ESTUDIOS DEL SUSTENTADOR PRINCIPAL			
Analfabeto o sin estudios	42,1	21,7	20,8
Primarios, EGB o FP-1	41,0	55,6	54,6
BUP, COU o FP-2	8,7	11,3	52,1
Diplomado Universitario	4,6	6,7	58,8
Estudios Superiores	3,7	4,6	50,6
CONDICIÓN SOCIOECONÓMICA DEL SUSTENTADOR PRINCIPAL			
Empresarios y direc. agrarios	3,6	5,1	56,8
Resto trabaj. y miem. coop. agrarias	6,5	6,0	36,8
Empresarios, profes. y trabaj. indep. no agrarios	8,3	12,1	59,0
Directores profes. y jefes administ. por cuenta ajena .	7,0	11,5	66,3
Resto personal servicios. y profes. FF.AA.	15,1	21,3	56,7
Contra maestres, operarios cualific. y miembros de coop. no agrarias	21,9	30,7	56,3
Operarios sin especialización, no agrarios	1,2	1,1	35,6
No clasificables	36,3	12,2	13,5
CATEGORÍA SOCIOPROFESIONAL DEL SUSTENTADOR PRINCIPAL			
Trabajadores manuales de la industria y servicios ...	19,0	28,0	59,2
Trabajadores no manuales de la industria y servicios	19,8	28,1	56,9
Autónomos de la industria y servicios	8,1	11,8	58,2
Trabajadores de la agricultura	9,6	10,4	43,5
Jubilados	27,3	9,6	14,1
Otros	16,2	12,2	30,5
TOTAL	100,0	100,0	40,2

circunscribimos a los hogares con algún miembro estudiando, esta participación se eleva al 84%. Entre los pensionistas se da una situación totalmente opuesta, ofreciendo un resultado congruente con lo que se obtenía al clasificar los hogares atendiendo a la principal fuente de ingresos. Curiosamente, es igual la proporción de parados entre los sustentadores principales de total de hogares y de los hogares con gasto en enseñanza.

Más de la mitad de los hogares tienen como sustentador principal una persona con estudios Primarios, EGB o Formación Profesional de primer grado. Al clasificar la población de acuerdo al nivel de estudios del miembro que presenta mayores ingresos, podemos resaltar como resultado más significativo el hecho de que, para los casos en que no disponga de estudios terminados, mientras que para el total de hogares esta situación se presenta en un 42% de los casos, esta cuota baja a la mitad entre los hogares con miembros estudiando. La explicación es bien simple: este nivel de estudios se presenta más frecuentemente en el grupo de edad de 65 y más años, y de entre los hogares que tienen como sustentador principal a una persona de este grupo, tan solo el 4% tienen algún miembro realizando estudios de primer o segundo grado (tabla 2).

De acuerdo con la condición socioeconómica del sustentador principal, observamos que la categoría más numerosa de hogares con miembros cursando estudios, con más del 30% de los casos, corresponde a contraamaestres, operarios cualificados y miembros de cooperativas no agrarias. A cierta distancia se sitúa el personal de servicios y profesionales de las Fuerzas Armadas, con exclusión de directores y jefes administrativos, que representa el 21%.

Trabajadores manuales y no manuales de la industria y los servicios, con un 28% en cada uno de los dos casos, son las categorías socioprofesionales que se dan con más frecuencia en el sustentador principal de hogares con gastos en enseñanza. Tan solo el 10,4% corresponde a trabajadores de la agricultura.

La consideración conjunta de las características analizadas en la tabla 2, nos informa que en el 75% de los casos el sustentador principal del hogar responde al siguiente perfil: es varón, de 30 a 64 años y ocupado; si además de estas tres características se considera el nivel de estudios, se observa en el 66% de los hogares con gastos en enseñanza que el sustentador principal es varón, de 30 a 64 años, ocupado y con estudios no universitarios.

Otra cuestión importante es si existen diferencias destacables entre los hogares con miembros inscritos en centros públicos o privados, se trata, pues, de identificar las características principales determinantes de la elección del centro escolar.

La característica «Principal fuente de ingresos del hogar» (ver tabla 3), no presenta grandes diferencias entre los hogares. Se observa un ligero aumento de la participación de la categoría «Por trabajo por cuenta ajena» en los centros privados, y por el contrario, un ligero descenso en la cuota que corresponde a alumnos pertenecientes a hogares con principal fuente de ingresos proveniente de trabajo por cuenta propia.

En cambio, la variable «nivel de ingresos del hogar», como era de esperar, presenta distribuciones muy dispares según sea la titularidad del centro educativo: en los

centros públicos se presentan mayores participaciones en los niveles inferiores, mientras, que en los de titularidad privada aumentan considerablemente las cuotas en los niveles superiores. El ejemplo más claro se produce en la categoría definida por ingresos globales del hogar superiores a 4 millones de pesetas, categoría que representa el

Tabla 3

DISTRIBUCIÓN DE LOS ALUMNOS SEGÚN LA TITULARIDAD DEL CENTRO POR ALGUNAS CARACTERÍSTICAS DEL HOGAR Y DEL SUSTENTADOR PRINCIPAL

	Centros Públicos (%)	Centros Privados (%)
PRINCIPAL FUENTE DE INGRESOS DEL HOGAR		
Por trabajo por cuenta ajena	72,5	79,4
Por trabajo por cuenta propia	20,0	13,9
Por pensión de jubilación y otros subsidios	6,9	6,6
Otros	0,8	—
NIVEL DE INGRESOS DEL HOGAR		
Menos de 1.000.000	13,0	6,4
De 1.000.000 a 2.000.000	47,4	31,6
De 2.000.000 a 4.000.000	34,1	42,0
Más de 4.000.000	5,5	20,0
RELACIÓN CON LA ACTIVIDAD DEL SUSTENTADOR PRINCIPAL		
Ocupados	82,1	93,9
Parados	8,2	2,3
Pensionistas	9,2	2,9
Otros inactivos	0,2	0,9
NIVEL DE ESTUDIOS DEL SUSTENTADOR PRINCIPAL		
Analfabeto o sin estudios	27,6	7,2
Primarios, EGB o FP-1	56,3	60,4
BUP, COU o FP-2	8,3	11,7
Diplomado Universitario	6,7	5,4
Estudios Superiores	1,1	15,4
TOTAL	100,0	100,0

20% de los alumnos en los centros privados, mientras que esta participación se reduce a un exiguo 5,5% en los centros públicos.

La participación de los alumnos pertenecientes a hogares con sustentador principal parado o pensionista es mucho menor en los centros privados que en los públicos. Por contra, el 94% de los alumnos de centros privados proviene de hogares en los que la persona de mayores ingresos está ocupada.

Por nivel de estudios, viene a ocurrir que para niveles superiores es mucho más elevada la proporción de alumnos en centros privados que públicos, si bien esta tendencia se rompe con los diplomados universitarios.

4. EL GASTO EN ENSEÑANZA DE PRIMER Y SEGUNDO GRADO

En la tabla 4 se presentan los gastos totales en enseñanza en el curso 90/91, detallados siempre que es posible (es decir, sólo para los gastos regulares y de inscripción) atendiendo a la titularidad del centro.

Si nos fijamos en la distribución del gasto por niveles educativos, vemos que más del 50% del gasto en enseñanza de los niveles considerados corresponde a EGB, mientras que la Educación Preescolar representa el 17,6%. No obstante, esta distribución de por sí no nos indica nada, pues, lógicamente, el gasto en cada uno de los grupos depende de manera fundamental del número de alumnos que lo componen. Así, en el

Tabla 4
GASTOS TOTALES EN ENSEÑANZA EN EL CURSO 90/91

NIVELES EDUCATIVOS	TOTAL	GASTOS REGULARES Y DE INSCRIPCIÓN		GASTOS COMUNES
		Centros Públicos	Centros Privados	
Educación Preescolar	883.970.113	45.095.056	511.127.639	327.747.418
EGB (Ciclos inicial y medio) . . .	1.650.218.332	—	207.069.918	1.443.148.414
EGB (Ciclo superior)	907.278.279	—	36.464.009	870.814.270
BUP	737.979.139	—	348.262.862	389.716.277
COU, Selectividad y pruebas de acceso a mayores de 25 años . . .	389.819.365	10.482.348	182.530.970	196.806.047
FP 1 ^{er} grado	258.203.511	—	7.802.463	250.401.048
FP 2 ^o grado	204.008.780	—	11.429.583	192.579.197
TOTAL	5.031.477.518	55.577.404	1.304.687.444	3.671.212.670

curso 90/91, los alumnos cursando estudios de EGB suponen el 60% de la cifra total, lo que indicaría que, en términos relativos, los gastos en este nivel de enseñanza son inferiores a los que incurren los hogares con hijos matriculados en los demás niveles educativos.

En efecto, en la tabla 5 puede comprobarse que, salvo en el primer grado de Formación Profesional, los gastos educativos por alumno de los hogares murcianos son, respecto a los demás niveles, con diferencia, los más bajos, 18.287 pesetas/alumno en el ciclo superior de la EGB o 23.470 pesetas de media por alumno en los ciclos inicial y medio, frente a 54.034 pesetas de COU, Selectividad y Pruebas de Acceso de mayores de 25 años. Un desglose del gasto nos informa que la razón fundamental de estas diferencias de gasto por alumno radica en la rúbrica de gastos regulares y de inscripción,

Tabla 5

GASTOS POR ALUMNO EN EL CURSO 95/96 POR NIVELES EDUCATIVOS

TIPO DE GASTO	IMPORTE
EDUCACIÓN PREESCOLAR	35.621
Gastos regulares y de inscripción	22.414
Gastos comunes	13.207
EGB (Ciclos inicial y medio)	23.470
Gastos regulares y de inscripción	2.945
Gastos comunes	20.525
EGB (Ciclo superior)	18.287
Gastos regulares y de inscripción	735
Gastos comunes	17.552
BUP	29.808
Gastos regulares y de inscripción	14.067
Gastos comunes	15.741
COU, SELECTIVIDAD Y PRUEBAS ACCESO MAYORES DE 25 AÑOS	54.034
Gastos regulares y de inscripción	26.754
Gastos comunes	27.280
FORMACIÓN PROFESIONAL 1^{er} GRADO	19.213
Gastos regulares y de inscripción	581
Gastos comunes	18.633
FORMACIÓN PROFESIONAL 2^o GRADO	24.186
Gastos regulares y de inscripción	1.355
Gastos comunes	22.831

apenas significativa en los dos primeros niveles (735 pesetas/alumno y 2.945 pesetas/alumno, respectivamente) frente a 26.754 pesetas que, como media, tiene que desembolsar cada alumno matriculado en COU, Selectividad o Pruebas de Acceso.

Sin embargo, de lo anterior no puede inferirse que exista una relación creciente entre gasto por alumno y nivel educativo. En este sentido, cabe destacar que, salvo en el nivel más próximo a la Universidad, que, como hemos dicho, es el que resulta más oneroso para los hogares murcianos, el nivel inicial de estudios, el de educación preescolar, con 35.621 pesetas, ocupa el segundo lugar en la escala de gastos medios por alumno. Como cabía esperar, son los gastos regulares y de inscripción la causa explicativa de estas diferencias, ya que los hogares con hijos en preescolar no gozan de las ventajas de la gratuidad de la enseñanza percibidas por los de EGB y FP.

En las tablas 6 y 7 aparece un desglose por niveles educativos de los gastos por alumno según la titularidad, pública o privada, del centro escolar.

Las cifras son bastante esclarecedoras de las diferencias de gasto por alumno de los hogares con hijo estudiando en cada centro. Así, un hogar con hijos inscritos en un centro privado le supone 2.5 veces más de gasto en el curso 95/96 que si estuviera matriculado en un centro público. Si bien esta diferencia aparece en todos los niveles educativos, la discrepancia, que apenas se aprecia en los ciclos que configuran la Formación Profesional y la EGB, sí que es significativa en Educación Preescolar (6 veces más que en centros públicos), y en COU, Selectividad y pruebas de acceso (7.8 veces superior). Una vez más, la causa de estas diferencias de gasto residen en la voluntariedad, y por tanto la no gratuidad, de sendos niveles educativos.

5. COMPONENTES PRINCIPALES DEL GASTO EDUCATIVO DE LOS HOGARES

El análisis de regresión es un método estadístico que propone modelos para representar la relación estadística entre dos o más variables, de manera que se pueda predecir, o explicar, el valor que tomará la variable dependiente, o explicar a partir de los valores que han tomado las variables independientes. Se trata, pues, de cuantificar la influencia que ejercen las variables explicativas sobre una variable dependiente de carácter continuo. Cuando el modelo aproxima mediante una relación lineal la dependencia estadística entre la variable dependiente y las independientes se denomina de regresión lineal múltiple.

En este apartado analizamos la relación que existe entre la variable GEDUC (gasto en educación) y las variables explicativas (componentes del gasto en educación): GCOMED (gasto en comedor escolar), GESCOLAR (gasto en material escolar), GINSCRIP (gasto en inscripción), GLIBROS (gasto en libros de texto), GOTRAS (gasto en otras actividades educativas), GPADRES (gasto en asociaciones de padres), GREGULAR (gastos regulares), GSEGURO (gastos de seguro escolar), GTRANS (gasto en transporte escolar) y GVARIOS (gastos varios). Así mismo, una vez cuantificada la relación, estudiaremos la contribución que cada variable independiente tiene en la variable GEDUC, claro está, para una familia de la Comunidad Autónoma de la Región de Murcia.

Tabla 6
GASTOS POR ALUMNO POR NIVELES EN LOS CENTROS PÚBLICOS
EN EL CURSO 95/96

NIVELES EDUCATIVOS	TOTAL	Gastos regulares y de inscripción	Gastos comunes
Educación Preescolar	15.641	2.433	13.207
EGB (Ciclos inicial y medio)	20.525	—	20.525
EGB (Ciclo superior)	17.552	—	17.552
BUP	15.741	—	15.741
COU, Selectiv. y pruebas acceso > 25 años ...	29.236	1.956	27.280
FP 1 ^{er} grado	18.633	—	18.633
FP 2 ^o grado	22.831	—	22.831
TOTAL	18.854	367	18.487

Tabla 7
GASTOS POR ALUMNO POR NIVELES EN LOS CENTROS PRIVADOS
EN EL CURSO 95/96

NIVELES EDUCATIVOS	TOTAL	Gastos regulares y de inscripción	Gastos comunes
Educación Preescolar	94.540	81.332	13.207
EGB (Ciclos inicial y medio)	31.755	11.230	20.525
EGB (Ciclo superior)	20.476	2.924	17.552
BUP	115.299	99.558	15.741
COU, Selectiv. y pruebas acceso > 25 años ...	228.954	201.674	27.280
FP 1 ^{er} grado	20.659	2.026	18.633
FP 2 ^o grado	35.493	12.662	22.831
TOTAL	46.634	28.148	18.487

Las características descriptivas de cada variable están en la tabla 8.

El análisis de la matriz de correlaciones para las once variables en estudio descarta problemas de multicolinealidad, pues las correlaciones entre las variables independientes son muy bajas (inferiores en todo caso a 0,2, y en muchos casos del orden de las centésimas). Así mismo, se tiene que las correlaciones entre la variable dependiente GEDUC y las independientes GRÉGULAR, GESCOLAR y GLIBROS son respectivamente 0.674, 0.525 y 0.480.

Tabla 8

CARACTERÍSTICAS DE LAS VARIABLES CONSIDERADAS EN EL ESTUDIO

VARIABLES	Media	Desviación típica	Valor mínimo	Valor máximo
GEDUC	53.155	68.679	14	425.180
GCOMED	347	3.910	0	54.600
GESCOLAR	9.100	27.106	0	160.680
GINSCRIP	1371	4.602	0	40.000
GLIBROS	16.930	28.121	0	120.000
GOTRAS	5.636	16.641	0	128.000
GPADRES	2.133	5.191	0	36.000
GREGULAR	12.512	43.469	0	392.925
GSEGURO	187	720	0	8.000
GTRANS	4.871	17.503	0	117.000
GVARIOS	66	650	0	9.000

Descartada la presencia de estas complicaciones estadísticas, se han calculado las ecuaciones de la regresión utilizando el método «stepwise» (pasos sucesivos). Así, en cada paso, se van obteniendo para las variables que hay en ese momento en la ecuación de regresión, entre otros, los coeficientes de regresión, la desviación típica de los mismos, el coeficiente de determinación y los coeficientes de regresión de las variables estandarizadas (beta). Para las variables que no están en ese momento en la ecuación aparecen los coeficientes denominados «beta in», que son los coeficientes de regresión que tendría cada variable si entrase en la ecuación en el paso siguiente. Al incorporar una nueva variable en la ecuación cambian los coeficientes de regresión de las variables que ya estaban en la misma y se vuelve a calcular el coeficiente de determinación.

Los principales resultados obtenidos se recogen en la tabla 9.

Tabla 9
PRINCIPALES RESULTADOS DE LA REGRESIÓN LINEAL MÚLTIPLE

PASO	VARIABLES	R	Beta In
1	GREGULAR	0,6739	0,6739
2	GESCOLAR	0,8530	0,5229
3	GLIBROS	0,9254	0,3637
4	GTRANS	0,9622	0,2662
5	GOTRAS	0,9935	0,2489
6	GPADRES	0,9965	0,0817
7	GINSCRIP	0,9983	0,0635
8	GCOMED	0,9999	0,0570
9	GSEGURO	1,0000	0,0101
10	GVARIOS	1,0000	0,0095

A la vista de estos resultados se puede concluir que el coeficiente de correlación múltiple (R) vale con las cuatro primeras variables independientes 0.9622 y, prácticamente lo mismo al añadir nuevas variables al modelo. La ecuación de regresión en este cuarto paso es:

$$\text{GEDUC} = 8287,1 + 1,016 \text{ GREGULAR} + 1,156 \text{ GESCOLAR} + 0,978 \text{ GLIBROS} + 1,044 \text{ GTRANS.}$$

Por tanto, el modelo con cuatro variables es más parsimonioso y explica prácticamente lo mismo que si se le añadiese más variables.

Otro aspecto interesante del análisis de regresión es analizar la importancia relativa de las distintas variables que intervienen en la predicción de la variable dependiente. Esta puede interesar en dos sentidos:

- Cuando la variable se utiliza individualmente. En este caso se puede responder analizando la matriz de correlaciones (Tabla 10) llegando a la siguiente ordenación:
 GREGULAR (0.674), GESCOLAR (0.525), GLIBROS (0.480), GINSKRIP (0.290), GOTRAS (0.288), GPADRES (0.273), GTRANS (0.2548), GSEGURO (0.095), GCOMED (0.061) y GVARIOS (-0.032)
- Cuando la variable se utiliza junto a otras variables. Ahora se utilizan los coeficientes beta o el incremento de R al incluir una nueva variable. En nuestro estudio y utilizando cualquiera de los procedimientos anteriores resulta la siguiente clasificación en orden decreciente de contribución: GREGULAR, GESCOLAR, GLIBROS, GTRANS, GOTRAS, GPADRES, GINSKRIP, GCOMED, GSEGURO y GVARIOS.

6. CONCLUSIONES

Las conclusiones más significativas que podemos extraer de este trabajo son las siguientes:

1. Se ha desarrollado una metodología general que permite, a partir de la información contenida en la Encuesta de Presupuestos Familiares, estudiar la distribución del gasto en educación por hogar a nivel de Comunidad Autónoma.
2. Más de la mitad de los gastos educativos no universitarios de los hogares murcianos (el 67,6%) se destina a los niveles de Preescolar y EGB. Sin embargo, como es obvio, ello no quiere decir que el gasto medio por alumno sea más elevado, precisamente, en estos niveles. Muy al contrario, la información estadística manejada pone de manifiesto que los gastos por alumno de los dos ciclos de EGB son, con diferencia, los más bajos de cuantos niveles aquí se consideran. Un análisis detallado de las rúbricas de gasto en enseñanza señala como causa explicativa a los gastos regulares y de inscripción, al no gozar de las ventajas que la obligatoriedad, y subsiguiente gratuidad, de la enseñanza representan para los hogares con niños en Preescolar o COU, selectividad y pruebas de acceso.
3. Las diferencias de gasto medio por alumno entre los centros públicos y privados son bastante importantes. Un hogar con hijos matriculados en un centro privado incurre en un gasto 2,5 veces más elevado que si estuvieran en un centro público, siendo destacable la diferencia de gasto en preescolar (6 veces más) y en COU, selectividad y pruebas de acceso (7,8 veces más).
4. La aplicación del método de regresión de pasos sucesivos («stepwise») nos ha permitido identificar las componentes del gasto que contribuyen en mayor medida a explicar la variabilidad del gasto en educación de los hogares de la Región de Murcia. Se concluye que con cuatro variables se explica prácticamente lo mismo que si considerásemos el conjunto total de variables.

7. BIBLIOGRAFÍA

- Aranda, J. y Gómez, J. (1996). *Fundamentos de Estadística para Economía y Administración de Empresas*. P.P.U. Barcelona.
- Centro Regional de Estadística de Murcia (1995). *Anuario Estadístico de la Región de Murcia*. Comunidad Autónoma de la Región de Murcia.
- Gómez, J., Lafuente, M. y Losa, A. (1997). *Gasto en educación no universitaria en la Región de Murcia. Un estudio sobre la influencia de las características familiares en el nivel del gasto educativo y su asignación público-privada*. Actas de la XXIII Reunión de Estudios Regionales. Valencia.
- Grao, J. y Ipiña, A. (1996). *Economía de la Educación. Temas de estudio e investigación*. Estudios y Documentos nº 22. Departamento de Educación, Universidades e Investigación. Gobierno Vasco.
- Instituto Nacional de Estadística. *Encuesta de Presupuestos Familiares 90-91*.
- Instituto Nacional de Estadística. *Índice de precios al consumo*.

Johnes, G. (1993). *The economics of education*. Macmillan.

Oroval, E. (1995). *Planificación, evaluación y financiación de sistemas educativos*. Civitas.

Palafox, J., Mora, J.G. y Pérez, F. (1995). *Capital humano, Educación y Empleo*. Fundación Bancaja, Valencia.

Uriel, E., Moltó, M.L., Pérez, F., Aldás, J. y Cucarella, V. (1997). *Las cuentas de la educación en España y sus comunidades autónomas: 1980-1992*. Colección Economía Española, vol. 4. Fundación Argentaria. Madrid.

ANÁLISIS DE ESTRUCTURAS DE COVARIANZA EN EL ESTUDIO DE LA COMPETENCIA DOCENTE DEL PROFESOR UNIVERSITARIO

*José Manuel García Ramos**

Universidad Complutense de Madrid

RESUMEN

El presente artículo presenta los resultados principales de un estudio realizado utilizando Análisis de Estructuras de Covarianza (Modelos Estructurales y de Medida) como metodología de análisis y validación en la aplicación de un modelo de Evaluación Formativa del Profesorado Universitario en un Centro Adscrito a la Universidad Complutense de Madrid.

Una vez definidos los modelos alternativos y utilizando LISREL se llegó a la definición de varios modelos causales que representan adecuadamente relaciones entre las dimensiones definidas como explicativas de la actividad docente en el aula.

ABSTRACT

The present article, presents the principal results of a study made, by using the covarianza structure of Analysis (Structural Models an Measures) as methodology of analysis and validation, at the application of a model of Formative Evaluation of the University Professor, at a Institution, linked to the University Complutense of Madrid.

One the alternative models are defined, using LISREL, a definition was reached of various casual models, that adequately represent, relations between definite and explicativa academic activities.

* José Manuel García Ramos es Catedrático de Métodos de Investigación y Diagnóstico en Educación de la Universidad Complutense de Madrid y Director del Centro Universitario Francisco de Vitoria, adscrito a dicha universidad. Especialista en temas de Investigación Educativa y Evaluación ha publicado 5 libros y numerosos artículos sobre temas relacionados con los Estilos Cognitivos, Métodos de Investigación Educativa y Evaluación (de Instituciones Educativas y del Rendimiento del Alumno). Fue premio Extraordinario de Licenciatura y Doctorado. Primer Premio Nacional de Investigación Educativa en 1986 (MEC-CIDE) y Tercer Premio Nacional de Investigación Educativa en 1996 (MEC-CIDE).

INTRODUCCIÓN

En el contexto de la Educación Superior, la necesidad de establecer Teorías y Constructos que guíen la acción educativa no es menor que la que se precisa en otros niveles educativos, máxime cuando en nuestros días la educación superior se ha universalizado y los recursos para su mejora no son ilimitados. Conectamos así con el tema de la Calidad de las Instituciones de Educación Superior y con su correlato natural, la Evaluación de la Calidad de las Instituciones Universitarias.

No nos vamos a extender ahora en desarrollar estos temas, sobre los cuales han corrido y siguen corriendo ríos de tinta. Baste como justificación y como punto de partida la aceptación común de la necesidad de realizar estudios e investigaciones que faciliten la consolidación de un cuerpo teórico que nos permita guiar, de la mejor manera posible, nuestra intervención educativa en contextos universitarios. Por otra parte, a ello habrá que añadir la necesidad de Evaluación de los sistemas o modelos de calidad universitaria que se definan, en orden a su mejora permanente (evaluación formativa) o bien, a la clasificación de las Instituciones de Educación Superior, a la toma de decisiones Institucional interna, a la asignación de Recursos Públicos o Privados a las Instituciones o Centros Educativos de nivel superior, ... Ambas preocupaciones, Investigación sobre... y Evaluación de... son importantes y reflejan necesidades distintas pero complementarias (García Ramos, 1997a y 1997b).

El objetivo de este **trabajo de investigación** es contribuir al estudio de un constructo pedagógico vinculado a la acción del profesor: **el constructo «Competencia Docente»**, contextualizado **en el nivel universitario**. Y ello, como contribución pequeña, pero quizás relevante, al establecimiento de modelos de acción docente eficaz en el nivel universitario. Ciertamente, no desconocemos lo criticable de esta opción, ni los resultados de la investigación que durante muchos años se ha realizado sobre la «eficacia docente» del profesor. Quizás estos resultados no hayan sido todo lo valiosos que hubiera sido deseable. Quizás no se utilizó la metodología de investigación adecuada, básicamente correlacional. Quizás se partía de un constructo estático, no bien contextualizado y con pretensiones excesivas de generalización de los resultados... No lo sé.

Ciertamente, habrá que definir con la mayor precisión posible estos sistemas de Calidad Educativa Universitaria. Pero habrá que hacerlo dinámicamente, probando modelos teóricos, formas o modelos de acción docente... y observando resultados, entendidos aquí no solamente como aprendizajes individuales e inmediatos, sino como Productos educativos valiosos (individuales y sociales, mediatos e inmediatos) para la sociedad en la que dicha acción educativa se inscribe (satisfacción de la demanda empresarial, cambio de actitud o de valores, calidad de los conocimientos adquiridos, capacidad para aplicar esos conocimientos adquiridos en la realidad socio-profesional-laboral, mejora de la calidad de vida general de la sociedad, etc...).

No me cabe duda, en este contexto, de que la acción del profesor en el aula y fuera del aula, su acción docente, es uno de los elementos que más directamente se vincula a todos estos elementos de calidad educativa mencionados. Y no le cabe duda, tampoco, a la sociedad y a sus responsables políticos, tanto a nivel estatal como regional, local o institucional. Cualquier política educativa y cualquier proyecto educativo pasa

por sus agentes principales: los docentes. De ahí se deriva, entre otras razones, que al valorar la calidad educativa de un sistema o de una institución educativa se comienza casi siempre por la evaluación de la acción docente. La universidad actual está llena de ejemplos que ratifican este hecho.

La muestra base de validación del presente trabajo la constituyen 4.109 cuestionarios CEDA, correspondientes a las valoraciones que los alumnos del Centro Universitario Francisco de Vitoria (adscrito a la Universidad Complutense) realizaron de sus profesores en Abril de 1995 (curso 94-95). En cada aplicación se superó el 80% de los alumnos presentes en la clase, el día de la valoración. El cuadro siguiente refleja la estructura de la muestra.

Tabla 1
MUESTRA POR CARRERAS Y SEXO

Carrera	Sexo		Total %
	Varón 1	Mujer 2	
A. Derecho	516	371	887 21.8
B. Economía	277	127	404 9.9
C. Periodismo	447	896	1.343 33.0
D. Empresariales	890	548	1438 35.3
Total	2.131	1.942	4.073
%	52.3	47.7	100.0

I. ANÁLISIS DE ESTRUCTURAS DE COVARIANZA (MODELOS ESTRUCTURALES Y DE MEDIDA) CON EL CEDA

Al objeto de completar el Análisis Estructural del CEDA, en nuestra aproximación al Constructo «Competencia Docente» del profesor universitario, formulamos varios modelos causales alternativos que implicaban, no sólo un modelo de medida, sino además un modelo estructural de relaciones entre las variables latentes (ver trabajos anteriores, García Ramos 1997a y 1997b).

En este caso tenemos modelos con cuatro tipos de variables, que en la nomenclatura del programa LISREL se definen como:

- X: Exógenas Medidas (Ítems del CEDA).
- Y: Endógenas Medidas (Ítems del CEDA).
- K: Exógenas Latentes.
- E: Endógenas Latentes.

El objetivo es el mismo que en los A.F.C. (García Ramos, 1997b); probar el ajuste de los modelos teóricos a los datos sobre la base de hipótesis establecidas. En este caso la complejidad es mayor, tanto en la definición de los Modelos, como en los restantes pasos de cualquier estudio causal (identificación, estimación de parámetros y evaluación de modelo).

Básicamente, vamos a contrastar la plausibilidad de 3 modelos Estructurales y de Medida, que varían entre sí según la naturaleza que se define para algunas de las variables latentes, (endógenas o exógenas) y las relaciones que entre las variables latentes se establece.

En síntesis, estos modelos se definen de la siguiente manera:

Modelo EM-1: 6 variables latentes, 4 endógenas y 2 exógenas correlacionadas (tabla 2).

Se parte de la hipótesis de que la «**Competencia Docente**» (E3 y E4) es **explicada directamente por las variables** «**Interacción con el Grupo – Participación**» (E1) y «**Atención individual al alumno**» (E2) y directa e indirectamente (a través de E1 y E2) por las variables exógenas correlacionadas «**Planificación–organización**» (K1) y «**Evaluación–Exámenes**» (K2).

En este modelo, se considera a E1 y E2 (Interacción con grupo y Atención individual) mediacionales y más cercanas a la explicación de la variabilidad de la Competencia Docente. El papel de la «**Evaluación**» en este modelo es más «formal» que en los siguientes (exógena y no explicada por otras en el modelo).

Modelo EM-2: 5 variables latentes, 4 Endógenas y 1 Exógena (tabla 13).

Se parte de la hipótesis de que la **variable** «**Evaluación–Exámenes**» es **endógena y mediacional o moderadora de la influencia de los otros 3 factores o variables latentes** («Planificación, «Interacción con grupo» y «Atención individual», siendo sólo la primera de ellas exógena). En este modelo se supone que la variable latente «Evaluación» está más próxima que las demás al juicio del alumno sobre la Competencia Docente del Profesor.

Modelo EM-3: 5 variables latentes, 1 Endógena y 3 Exógenas (tabla 22).

Se parte también de la consideración de la **naturaleza mediacional de la variable** «**Evaluación– Exámenes**» (al igual que en EM-2), **pero considerando Exógenas a las otras tres variables explicativas**. Es un modelo menos parsimonioso que los anteriores y más próximo a un modelo predictivo. Por tanto, teóricamente más débil y menos aceptable.

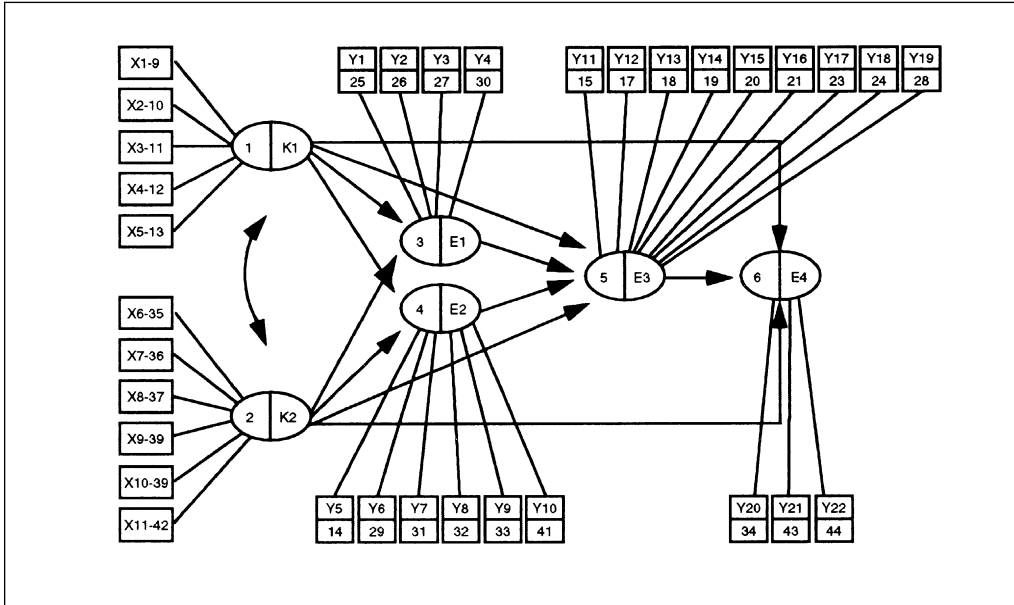
Sobre estos modelos, realizamos variaciones en función de los resultados obtenidos.

1.1. Modelo EM-I

a. Modelo General

El siguiente cuadro muestra las características del modelo:

Tabla 2
MODELO ESTRUCTURAL Y DE MEDIDA 1 (EM-1)



Donde tenemos:

- 11 variables empíricas exógenas (X)
- 22 variables empíricas endógenas (Y)
- 4 variables latentes exógenas (K)
- 4 variables latentes endógenas (E)

Siendo:

- K1 - Planificación/Organización
- K2 - Evaluación/Exámenes
- E1 - Interacción con el grupo/Participación
- E2 - Atención individual al alumno
- E3 - Competencia Docente
- E4 - Calidad Docente (Items criterio)

Nota: las variables empíricas (medias) aparecen definidas también por el nº de ítems del CEDA.

Este modelo, como puede verse, parte de la hipótesis de que los Factores de Planificación (K1) y Evaluación (K2), son variables exógenas latentes correlacionadas entre sí, que inciden en la Calidad docente (E4) y en la Competencia docente (E3) directa e indirectamente, a través de los factores Participación (E1) y Atención personal al alumno (E2). La definición de la estructura de medida de cada variable latente es la establecida en la Construcción del CEDA, que estaría básicamente contrastada en el modelo de A.F.C. nº 3 (García Ramos, 1997b). La definición de la estructura causal de relaciones entre las variables latentes parte de la consideración de que las variables programación (K1) y Evaluación (K2) son variables causales exó-

genas, no explicadas por ninguna otra variable latente y correlacionadas entre sí. Su influencia sobre la Competencia Docente (E3) y sobre la Calidad Docente (E4) se encontraría, en parte, mediatizada por factores de relación–motivación propios de la relación profesor–alumno que se establece en clase (E1) y fuera de la clase (E2).

Las variables latentes E3 y E4, en realidad son la misma dimensión de calidad–competencia docente, que se miden en el CEDA a través de Escalas diferentes: **analíticamente** (subescalas B y C del CEDA, implícitas en la dimensión E3) y **globalmente** (ítems–criterio 34, 43 y 44). En un análisis posterior eliminaremos la variable E4 (definida por ítems–criterio, que valoran globalmente la calidad docente) para trabajar sólo con ítems que aprecian analíticamente la conducta docente (los implicados en las variables latentes K1, K2, E1, E2 Y E3).

En el modelo EM–1 de la tabla anterior, se tomaron los siguientes ítems como referencia de medida de las variables latentes, fijando a 1.00 su parámetro estructural, coherentemente con los A.F.E y A.F.C. anteriores (García Ramos, 1997a y 1997b):

- K1 – Ítem 12 (Orden)
- K2 – Ítem 36 (Ajusteva)
- E1 – Ítem 26 (Motivaci)
- E2 – Ítem 33 (Atención)
- E3 – Ítem 18 (Resaltar)
- E4 – Ítem 44 (Comparat)

El número total de parámetros que debe calcular LISREL es de 81:

- a) 18 para los ítems de las 4 variables endógenas latentes (Matriz LAMBDA–Y).
- b) 9 para los ítems de las 2 variables exógenas latentes (Matriz LAMBDA–X).
- c) 8 correlaciones entre las variables K(2) y la E(4) ($4 \times 2 = 8$) (Matriz GAMMA).
- d) 3 correlaciones entre las variables K (Matriz PHi).
- e) 10 correlaciones entre las variables ETA (Matriz PSI).
- f) 22 errores de medida, de las variables Y (Matriz Theta Epsilon).
- g) 11 errores de medida, de las variables X (Matriz Theta Delta).

Ofrecemos a continuación las estimaciones de los parámetros de medida (las matrices LAMBDA–Y y LAMBDA–X) y las matrices Theta Eps. y Theta Delta.

Tabla 3
MATRICES LAMBDA-Y Y THETA EPSILON (MODELO EM-1)
(CORRELACIONES POLICÓRICAS)

Sub-Escalas CEDA	Items	Matriz LAMBDA-Y. Solución Final (Max. Lik.) FACTORES				THETA EPS. (S. Final)
		ETA 1	ETA 2	ETA 3	ETA 4	
• Interacción con el grupo (E-1)	25	.800				.466
	26	1.000				.166
	27	.621				.678
	30	.876				.359
• Atención personal (E-2)	14		.773			.518
	29		.842			.427
	31		.821			.455
	32		.939			.287
	33		1.000			.192
	41		.809			.471
• Competencia Docente (Analit) (E-3).	15			.963		.380
	17			.975		.364
	18			1.000		.331
	19			.910		.447
	20			.877		.486
	21			.990		.345
	23			.849		.518
	24			.955		.390
	28			.998		.334
• Ítems criterio (Competenciadocente global) (E-4)	34				.892	.281
	43				1.00	.091
	44				1.00	.096

Como puede observarse, las variables endógenas latentes son muy consistentes, especialmente ETA3 y ETA4 (Competencia docente, Analítica y Global). Los ítems 27 y 14 presentan bajos valores en sus parámetros estructurales y altos índices de error. También los ítems 23, 25, 20 y 41 presentan altos índices de error, aún cuando sus «saturaciones» son aceptables (son ítems «neutros», en cuanto que son ítems con sentido teórico, pero que no contribuyen al ajuste empírico del modelo).

Tabla 4
MATRICES LAMBDA-X Y THETA-DELTA (MODELO EM-1)
(CORRELACIONES POLICÓRICAS)

Subescalas CEDA	Items	Matriz Lambda-X-Solución Final (M.L.) FACTORES		Theta Delta (Sol. Final)
		KSI-1	KSI-2	
• Planificación – Organización. (K1)	9	.835		.508
	10	.691		.663
	11	.910		.416
	12	1.000		.295
	13	.964		.344
• Evaluación – Exámenes (K2)	35		.812	.487
	36		1.000	.222
	37		.627	.695
	38		.961	.282
	39		.915	.348
	42		.852	.435

Las variables exógenas latentes son consistentes. Los ítems 10 y 37 presentan baja «saturación» y alto índice de error. Los ítems 9, 35 y 42 pueden considerarse «neutros».

La tabla siguiente nos muestra los coeficientes de Determinación Total para las variables-X, las variables-Y y para las ecuaciones Estructurales del Modelo. Son valores muy satisfactorios en todos los casos:

Tabla 5
COEFICIENTE DE DETERMINACIÓN TOTAL DEL MODELO EM-1

- | |
|---|
| <ul style="list-style-type: none"> • Para las Variables-X = .984 • Para las Variables-Y = 1.000 • Para las Ecuaciones Estructurales = .859 |
|---|

Los valores de ajuste del modelo son desigualmente aceptables siendo el mejor RMSR; indican inadecuado ajuste del modelo y deben mejorarse, si fuera posible:

Tabla 6
ÍNDICES DE AJUSTE DEL MODELO EM-1
(CORRELACIONES POLICÓRICAS)

- Goodness of fit index (G.F.I.) = .759
- Adjusted goodness of fit index (A.G.F.I.) = .718
- Root Mean Square Residual (R.M.S.R.) = .067

Si en vez de correlaciones Policóricas, utilizamos correlaciones Pearson, los índices de ajuste del modelo mejoran sensiblemente:

Tabla 7
ÍNDICES DE AJUSTE DEL MODELO EM-1
(CORRELACIONES PEARSON)

- G.F.I. = .811
- A.G.F.I. = .779
- R.M.S.R. = .061

b) Modelo EM I.I. (Mejora del Modelo EM-1)

Si eliminamos las variables empíricas (x e y) que desajustan el modelo (27, 14, 23, 35, 37, 13, 10 y 41), manteniendo la misma definición de relaciones entre variables latentes (exógenas y endógenas) y utilizando correlaciones Pearson, los índices de ajuste del modelo mejoran sustancialmente, pudiéndose considerar satisfactorios en todos los casos.

Tabla 8
ÍNDICES DE AJUSTE DEL MODELO EM-1, ELIMINANDO ÍTEMS
(CORRELACIONES PEARSON)

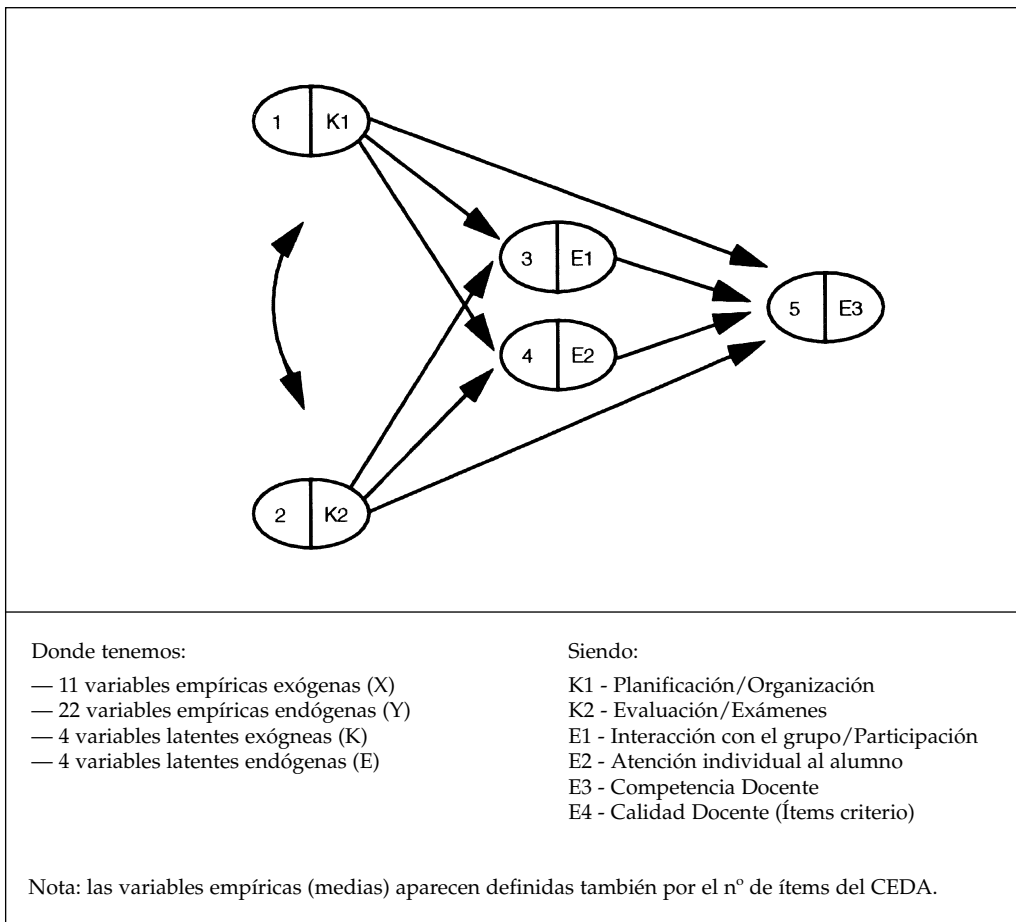
- G.F.I. = .882
- A.G.F.I. = .851
- R.M.S.R. = .050

c) Modelo EM-1.2 (Eliminando ETA-4 del Modelo EM-1)

Si eliminamos el factor E-4, por redundante y para homogenizar el modelo, utilizando sólo ítems que valoran analíticamente la Competencia Docente del profesor, los resultados apenas varían respecto a los análisis anteriores.

En este caso el modelo sería el siguiente (definiendo de la misma manera que en EM-1, cada variable latente):

Tabla 9
MODELO EM-1.2



Sólo mostramos los índices de ajuste del Modelo, pues las demás matrices que lo definen apenas varían respecto al modelo EM-1.

Tabla 10
ÍNDICES DE AJUSTE DEL MODELO EM-1.2
(CORRELACIONES POLICÓRICAS)

- G.F.I. = .762
- A.G.F.I. = .718
- R.M.S.R. = .078

Si utilizamos correlaciones Pearson, los índices de ajuste del modelo mejoran, pero no superan los valores obtenidos en el contraste del Modelo EM-1.

Tabla 11
ÍNDICES DE AJUSTE DEL MODELO EM-1.2
(CORRELACIONES PEARSON)

- G.F.I. = .811
- A.G.F.I. = .779
- R.M.S.R. = .061

Si eliminamos los ítems que desajustan el modelo (27, 14, 23, 35, 37, 13, 10 y 41) y utilizamos correlación Pearson obtenemos mejores índices de bondad de ajuste del Modelo.

Tabla 12
ÍNDICES DE AJUSTE DEL MODELO EM-1.2, ELIMINANDO ÍTEMS
(CORRELACIONES PEARSON)

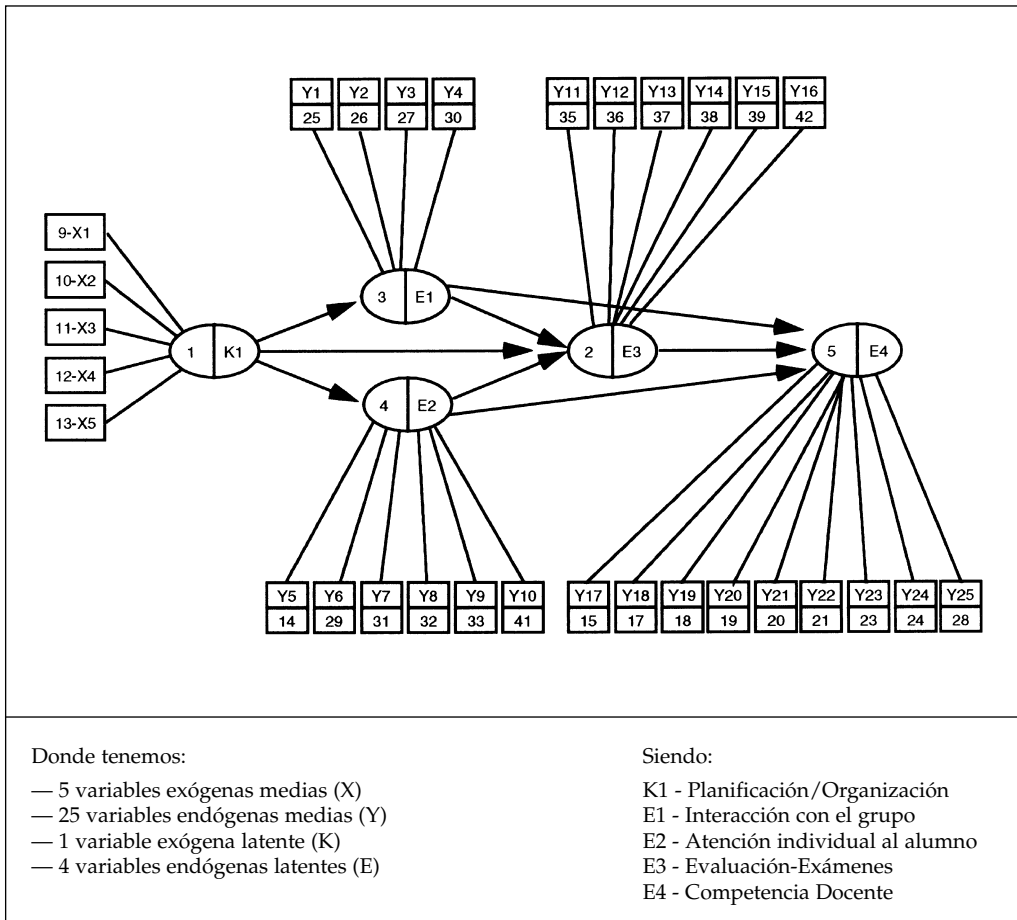
- G.F.I. = .877
- A.G.F.I. = .841
- R.M.S.R. = .053

I.2. Modelo EM-2

a) *Modelo General*

El siguiente cuadro nos muestra las características de este modelo, alternativa explicativa al modelo anterior (EM-1):

Tabla 13
MODELO ESTRUCTURAL Y DE MEDIA 2 (EM-2)



Este modelo parte de la hipótesis de que la variable latente «Evaluación-Exámenes» es una variable mediacional (a diferencia del Modelo EM-1, que la considera exógena) y por tanto endógena, que modera la influencia de los otros 3 factores

antecedentes en el modelo (K1, E1 y E2) sobre la valoración por el alumno de la Competencia Docente del Profesor Universitario (E4). La definición, en la estructura de medida de todas las variables latentes, es idéntica a la establecida en el modelo EM-1 y coherente con los resultados de los A.F.E. y A.F.C. anteriores.

En este caso, el número de parámetros a estimar por LISREL es de 74. Las estimaciones de los parámetros de medida son similares a las obtenidas en el modelo anterior, lo cual es lógico, dado que las variables latentes, como modelos de medida, han sido definidas de igual manera.

Tabla 14
 MATRICES LAMBDA Y Y THETA EPSILON (EM-2)

Sub-Escalas CEDA	Ítems	Matriz LAMBDA-Y. Solución Final (Max. Lik.) FACTORES				THETA EPS. (S. Final)
		ETA 1	ETA 2	ETA 3	ETA 4	ERROR
• Interacción con el grupo (E- 1)	25	.801				.465
	26	1.00				.165
	27	.621				.678
	30	.875				.361
• Atención personal (E-2)	14		.775			.513
	29		.842			.426
	31		.819			.457
	32		.936			.291
	33		1.000			.191
	41		.810			.469
• Evaluación - Exámenes (E3)	35			.814		.484
	36			1.000		.222
	37			.629		.692
	38			.960		.282
	39			.914		.349
	42			.849		.439
• Competencia Docente (Analit) (E-3).	15				.963	.378
	17				.970	.368
	18				1.000	.329
	19				.913	.440
	20				.884	.479
	21				.993	.338
	23				.833	.534
	24				.948	.396
	28				.995	.335

Tabla 15
MATRICES LAMBDA X Y THETA DELTA (EM-2)

Subescalas CEDA	Items	Matriz Lambda-X-Solución Final (M.L.)	Theta Delta (Sol. Final)
		KSI-1	ERROR
• Planificación- Organización (K1)	9	.837	.502
	10	.692	.660
	11	.911	.410
	12	1.000	.289
	13	.950	.358

Los Coeficientes de Determinación total son similares también a los del modelo EM-1, siendo algo inferior el valor correspondiente a las Ecuaciones Estructurales.

Tabla 16
ÍNDICES DE DETERMINACIÓN TOTAL DEL MODELO EM-2

- | |
|---|
| <ul style="list-style-type: none"> • Para las Variables-X = .878 • Para las Variables-Y = 1.000 • Para las Ecuaciones Estructurales = .788 |
|---|

Los índices de Bondad de ajuste del Modelo son muy similares a los del modelo EM-1. Indican bajo nivel de ajuste en las dos primeras estimaciones, siendo más aceptable R.M.S.R.

Tabla 17
*ÍNDICES DE BONDAD DE AJUSTE DEL MODELO EM-2 (CORRELACIONES
POLICÓRICAS)*

- | |
|---|
| <ul style="list-style-type: none"> • G.F.I. = .762 • A.G.F.I. = .717 • R.M.S.R. = .070 |
|---|

Si utilizamos correlaciones de Pearson, el nivel de ajuste mejora sensiblemente, al igual que ocurría en el modelo EM-1.

Tabla 18

ÍNDICES DE BONDAD DE AJUSTE DEL MODELO EM-2 (CORRELACIONES PEARSON)

- G.F.I. = .815
- A.G.F.I.= .780
- R.M.S.R.= .063

Los resultados no varían si se elimina la restricción de incorrelación entre las variables 1 y 5 y entre las variables 3 y 4.

b) Modelo EM-2.1. (Mejora del Modelo EM-2)

Si eliminamos los ítems que desajustan el modelo (27, 14, 23, 35, 37, 13, 10 y 41), utilizando correlaciones de Pearson y admitiendo la existencia de relación entre E1 y E2 y la influencia directa de la variable exógena latente (K1) sobre E4 (Competencia Docente), mejoramos los índices de ajuste generales del modelo (índices de bondad de ajuste) pero descienden otros indicadores de ajuste, tales como por ejemplo, los Coeficientes de Determinación de las Variables-X y especialmente, los correspondientes a las Ecuaciones Estructurales.

Tabla 19

COEFICIENTES DE DETERMINACIÓN TOTAL DEL MODELO EM-2.1

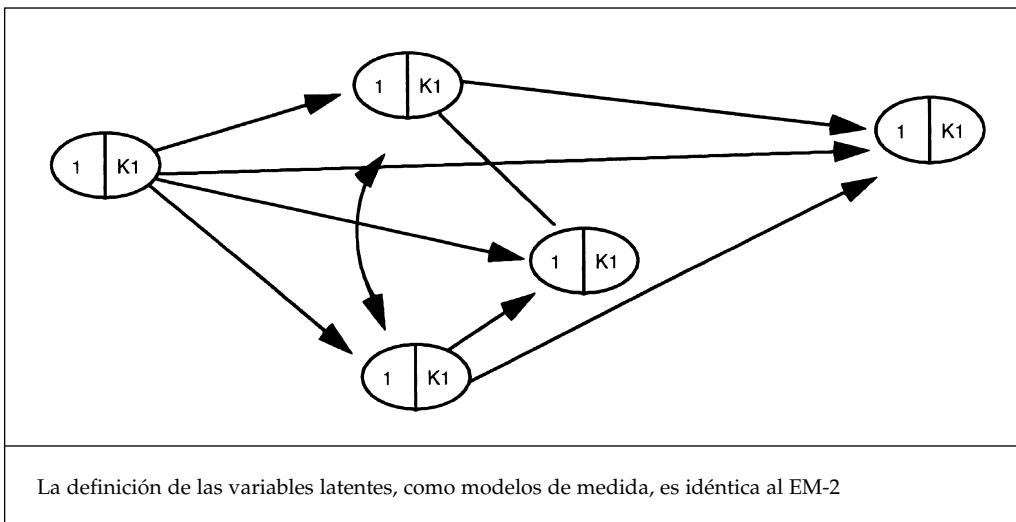


Tabla 20

COEFICIENTES DE DETERMINACIÓN TOTAL DEL MODELO EM-2.1.

- Para las Variables-X = .812
- Para las Variables-Y = .999
- Para las Ecuaciones Estructurales = .735

Tabla 21

ÍNDICES DE BONDAD DE AJUSTE DEL MODELO EM-2.1.

- G.F.I. = .877
- A.G.F.I. = .843
- R.M.S.R. = .053

I.3. Modelo EM-3

a) Modelo General

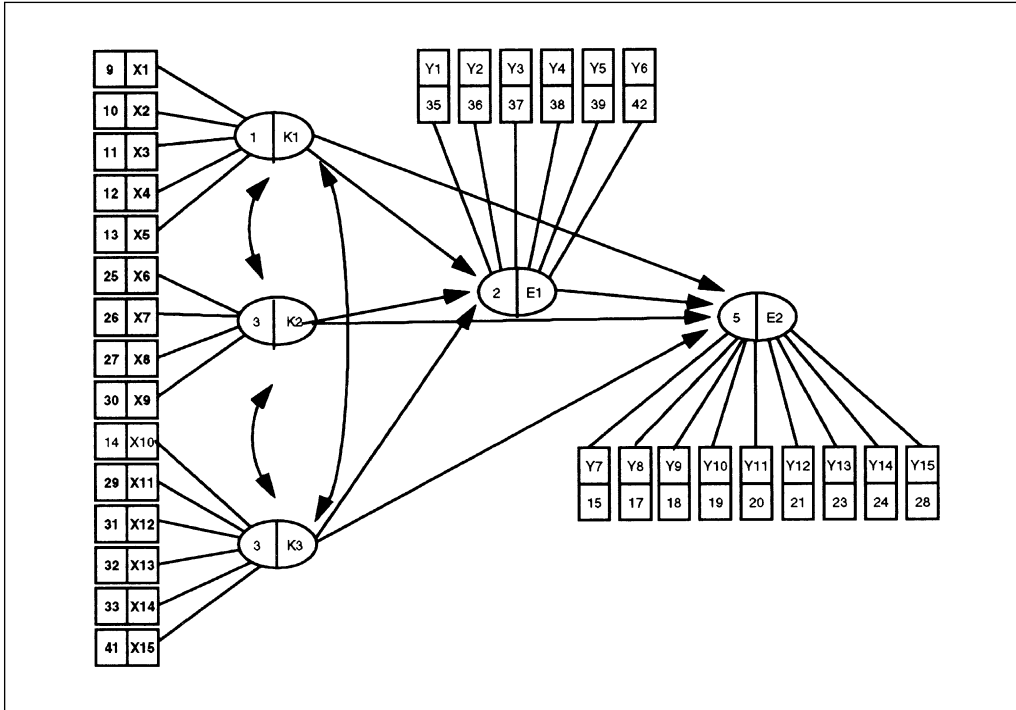
El siguiente modelo (Tabla 22) fue definido como alternativa explicativa a los dos anteriores (EM-1 y EM-2).

Este modelo, más conservador y cercano a un modelo predictivo, parte de la hipótesis de 3 variables exógenas latentes, correlacionadas entre sí, que inciden directa e indirectamente (a través del modelo evaluativo del docente) sobre la «Competencia Docente» del Profesor, según el juicio del alumno. La variable «Evaluación-Exámenes» se considera mediacional, como en el modelo EM-2. La definición de las variables latentes, como modelos de medida, es idéntica a la establecida en los modelos EM-1 y EM-2 y coherente con los A.F.E. y A.F.C. efectuados.

En este caso, el número de parámetros a estimar por LISREL es de 70.

Sin embargo, los Coeficientes de Determinación Total son mucho mejores, especialmente en lo que se refiere al Coeficiente de Determinación total para las ecuaciones estructurales, que alcanzan un muy satisfactorio valor (.909).

Tabla 22
 MODELO ESTRUCTURAL Y DE MEDIDA 3 (EM-3)



Donde tenemos:

- 15 variables exógenas medias (X)
- 15 variables endógenas medias (Y)
- 3 variables exógena latente (K)
- 2 variables endógenas latentes (E)

Siendo:

- K1 - Planificación/Organización
- E1 - Interacción con el grupo
- E2 - Atención individual al alumno
- E3 - Evaluación/Exámenes
- E4 - Competencia docente

La numeración de las variables latentes, para facilitar la comparación, es la misma que la definición en los otros dos modelos (EM1 y EM2).

Tabla 23
 COEFICIENTES DE DETERMINACIÓN TOTAL DEL MODELO EM-3

- Para las Variables-X = .997
- Para las Variables-Y = .988
- Para las Ecuaciones Estructurales = .909

Los índices de Bondad de Ajuste del Modelo son muy similares a los de los dos modelos anteriores. Indican bajo nivel de ajuste del modelo, excepto en el índice RMSR.

Tabla 24

ÍNDICES DE BONDAD DE AJUSTE DEL MODELO EM-3 (CORRELACIONES POLICÓRICAS)

- G.F.I. = .762
- A.G.F.I. = .719
- R.M.S.R. = .070

Si utilizamos correlaciones de Pearson, el nivel de ajuste del modelo mejora sensiblemente.

Tabla 25

ÍNDICES DE BONDAD DE AJUSTE DEL MODELO EM-3 (CORRELACIONES PEARSON)

- G.F.I. = .815
- A.G.F.I. = .782
- R.M.S.R. = .063

b) Modelo EM-3.1 (Mejora del Modelo EM-3)

Si eliminamos los ítems que desajustan el modelo (27, 14, 23, 35, 37, 13, 10 y 41) y utilizando correlaciones Pearson, mejoramos los indicadores de ajuste del modelo (que alcanzan valores satisfactorios), no disminuyendo sensiblemente los Coeficientes de Determinación Total.

Tabla 26

ÍNDICES DE BONDAD DE AJUSTE DEL MODELO EM-3.1 (CORRELACIONES PEARSON)

- G.F.I. = .877
- A.G.F.I. = .843
- R.M.S.R. = .053

I.4. Conclusiones del Análisis Comparativo de los distintos modelos Alternativos, Estructurales y de Medida, contrastados mediante análisis de Estructuras de Covarianza (Evaluación de modelos)

La tabla siguiente nos muestra, comparativamente, los principales indicadores de Ajuste de los 3 modelos estructurales y de medida, formulados como alternativas explicativas, con sus principales variantes.

Tabla 27
ANÁLISIS COMPARATIVO DEL NIVEL DE AJUSTE DE LOS MODELOS HIPOTÉTICOS ALTERNATIVOS, ESTRUCTURALES Y DE MEDIDA

			Indicadores de Bondad de Ajuste			Coeficientes de Determinación Total para las:			Referencias
Modelos	Submodelo	Matriz R	GFI	AGFI	RSMR	X	Y	Ec. Estr.	Anexo
EM-1	General (33 ítems)	Policóricas	.759	.718	0.67	.984	1.00	.859	VI.1
		Pearson	.811	.779	.061	.977	.999	.855	VI.2
	1.1. (25 ítems)	Pearson	.882	.851	.050	.970	.999	.818	VI.3
		Policóricas	.762	.718	.078	.984	.997	.851	VI.4
	1.2. (30 ítems)	Pearson	.877	.841	.053	.970	.995	.810	VI.5
EM-2	General (30 ítems)	Policóricas	.762	.717	.070	.878	1.00	.788	VI.6
		Pearson	.815	.780	.063	.868	.998	.785	—
	2.1 (22 ítems)	Pearson	.877	.843	.053	.812	.999	.735	VI.7
EM-3	General (30 ítems)	Policóricas	.762	.719	.070	.997	.988	.909	VI.8
		Pearson	.815	.782	.063	.996	.984	.906	VI.9
	3.1 (22 ítems)	Pearson	.877	.843	.053	.994	.983	.900	VI.10

Dado el diferente número de ítems incluidos en los distintos modelos, la comparación entre indicadores no es obvia, a primera vista. Por otra parte, la complejidad de los modelos es distinta, al incluirse en ellos las variables latentes con diferente caracterización (algunas variables actúan como exógenas en unos modelos y como endógenas en otros).

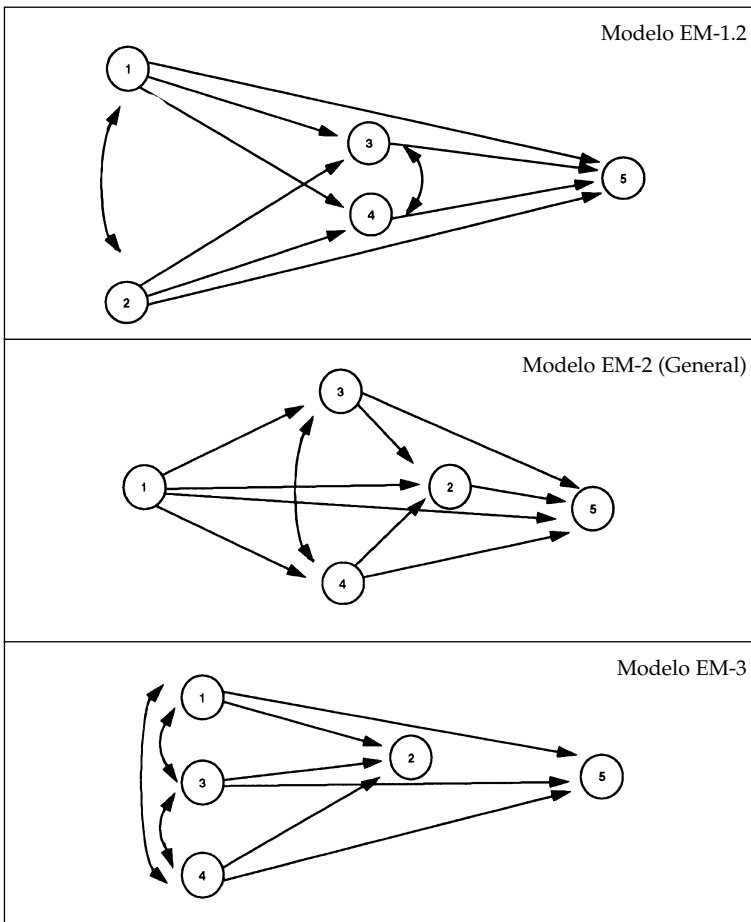
Sí parece posible, sin embargo, comparar legítimamente los modelos EM-1.2, EM-2 General y EM-3 General, pues los 3 responden a Modelos con el mismo número de ítems, en donde estos (los ítems) definen de la misma manera a los mismas variable latentes (idénticos modelos de medida de las variable latentes implicadas). En todos ellos se trabaja con las siguientes variables latentes:

1. Planificación–Organización.
2. Evaluación–Exámenes
3. Interacción con el grupo de clase
4. Atención individualizada al alumno
5. Competencia Docente del Profesor (valoración analítica).

El cuadro siguiente permite resumir la naturaleza de las relaciones, establecida entre las variables, en los distintos modelos:

Tabla 28

**MODELOS ESTRUCTURALES Y DE MEDIDA CON 30 ÍTEMS E IDÉNTICO
MODELO DE MEDIDA DE LAS VARIABLES LATENTES**



Si nos fijamos en los indicadores de Bondad de Ajuste de los 3 modelos, no hay apenas diferencias cuando trabajamos con la matriz de correlaciones policóricas, siendo mayor la diferencia, a favor del Modelo EM-1.2 cuando comparamos estos indicadores sobre la base de los análisis realizados con correlaciones Pearson. Si observamos los coeficientes de Determinación Total, vemos que el peor ajuste se da en el modelo EM-2 (bien es cierto que sólo hay una variable exógena), siendo mejores las estimaciones en EM-1.2 y mucho mejores las estimaciones en EM-3, especialmente para los Coeficientes de Determinación correspondientes a las Ecuaciones Estructurales, que alcanzan valores muy satisfactorios (alrededor de .91). Por tanto, parece descartable el Modelo EM-2, ante los mejores ajustes obtenidos en EM-1.2 y en EM.3. De estos dos modelos sólo podemos optar sobre la base de la Teoría, pues cada uno de ellos presenta cualidades empíricas distintas favorables:

- El Modelo EM-1, con 33 ítems es similar en sus valores de Bondad de Ajuste al Modelo EM-3, con 30 ítems, a pesar del «handicap» que supone la existencia de más ítems en el modelo y 1 variable latente más (Competencia Docente, valorada globalmente). El modelo EM-1.1 (idéntico al EM-1, pero eliminando ítems que desajustan el modelo) con 25 ítems alcanza valores de Bondad de Ajuste superiores al EM-3 (30 ítems) y al EM-3.1 (22 ítems).
- Por el contrario, el Modelo EM-3 y el EM-3.1, obtienen mejores indicadores en el Coeficiente de Determinación Total para las ecuaciones estructurales. Bien es verdad, que estos modelos, más convencionales y simples, se asemejan más a un modelo predictivo, con la única salvedad de la acción mediacional de la variable latente 2 (Evaluación-Exámenes).
- Desde el punto de vista teórico, es preferible el modelo EM-1, pues sitúa a las variables latentes 3 y 4, vinculadas a procesos de relación profesor-alumno, como mediacionales y explicativas, más próximas a la variable «Competencia Docente». Por otro lado, las variables 1 y 2 (Programación-Organización y Evaluación-Exámenes) son considerados más formalmente como variables exógenas, cuya influencia sobre la «Competencia Docente» es directa, pero también indirecta, a través de las variables 3 y 4 («Interacción el grupo» y «Atención personalizada»).

Por otra parte, y considerando complejo el Constructo «Competencia Docente» del Profesor Universitario, tal como lo valoran los alumnos, es preferible un modelo explicativo más complejo (el EM-1) y por tanto, más parsimonioso, más plausible como explicación de la estructura valorativa del estudiante, en su juicio sobre la Competencia Docente del Profesor Universitario.

En todo caso, se necesita más investigación para determinar o confirmar por otras vías, la causalidad que se establece entre las dimensiones implicadas en la valoración y juicio sobre la Competencia Docente del Profesor Universitario. Especialmente útiles podrían ser estudios por submuestras (carreras, tipos de asignaturas, etc...) en donde los diferentes modelos pueden ser diversamente plausibles.

2. CONCLUSIONES, INTERPRETACIÓN Y DISCUSIÓN DE RESULTADOS

Introducción

Tratar de identificar las dimensiones que conforman el perfil de un buen docente universitario, ha sido el trabajo de numerosos estudios e investigaciones en los últimos 30 años. Si revisáramos estos estudios veríamos que hay ciertos puntos de acuerdo sustancial sobre estas dimensiones: Trent y Cohen, 1973; Marsh, 1982 y 1987; Ramsden y Entwistle, 1981; Jornet y Suárez, 1988; Tourón, 1989; Entwistle y Tait, 1990; Ramsden, 1991; Dahllöf, 1990; Tejedor, 1993; Rodríguez Espinar, 1993; Cajide, 1994; etc...

A modo de comparación y como ejemplo veamos en el siguiente cuadro (García Ramos, 1997a) las principales dimensiones implícitas en algunos de los instrumentos utilizados, que han sido suficientemente probados y validados:

Tabla 49

DIMENSIONES IMPLICADAS EN DIFERENTES INSTRUMENTOS DE EVALUACIÓN DOCENTE (POR EL ALUMNO)

Dimensiones comunes	SEEQ	Endeavour	CEQ	E.E.P.A.	CEDA - MOD.
Trent y Cohen (1973)	Marsh (1982 y 87)	Frey, Leonard y Beatty (1975)	Ramsden (1991)	Tejedor (1993)	García Ramos (1996)
1. Estructura y claridad en la explicación. 2. Positiva actitud hacia la discusión en clase y la exposición de puntos diferentes. 3. Estímulo del interés, motivación y reflexión del alumno. 4. Prestar atención individual al alumno. 5. Entusiasmo.	1. Aprendizaje-Valor. 2. Interacción con el grupo. 3. Raport individual. 4. Exámenes- Calificaciones. 5. Carga de trabajo - Dificultad. 6. Organización-Claridad. 7. Entusiasmo. 8. Amplitud de enfoque. 9. Trabajo extraclase-Lecturas.	1. Logros de los alumnos. 2. Discusión en clase. 3. Atención personal. 4. Calificaciones. 5. Carga de trabajo. 6. Claridad de presentación. 7. Organización-Planificación.	1. Enseñanza adecuada. Correcta orientación sobre la marcha del aprendizaje del alumno. 2. Claridad en los objetivos del programa propio y en lo que se espera de uno. 3. Adecuada carga de trabajo. 4. Adecuada evaluación- calificación del aprendizaje. 5. Estímulo y motivación al alumno en su responsabilidad (A los estudiantes se les dan muchas opciones, en el trabajo que realizan).	1. Cumplimiento de obligaciones. 2. Calidad y desarrollo del programa. 3. Dominio de la asignatura. 4. Interacción con los alumnos. 5. Recursos utilizados y prácticas. 6. Exámenes.	1. Programación-Organización de la Enseñanza. 2. Dominio de contenidos-claridad expositiva. 3. Motivación de aprendizaje - Incremento de interés del alumno. 4. Interacción con el grupo de clase. 5. Atención individual al alumno. 6. Evaluación - Exámenes.

Las diferencias entre las distintas dimensiones definidas en los instrumentos, a mi juicio, tienen más que ver con formas personales de concebir la enseñanza universitaria, vinculadas al contexto en que están siendo definidas, que con otros factores.

De alguna manera, estos componentes están presentes (de un modo particular ciertamente) en los instrumentos utilizados en nuestra investigación.

2.1. Valoración de la Estructura de relaciones entre las dimensiones que definen el Constructo Competencia Docente del Profesor Universitario, medido a través del CEDA

El presente trabajo nos permitió valorar comparativamente varios modelos hipotéticos alternativos que establecían relaciones de causalidad diversa entre las variables latentes o dimensiones de medida estudiadas (tablas 2, 12 y 22, principalmente).

La tabla 27, permite comparar los principales indicadores de ajuste de los 3 modelos generales establecidos, así como de sus variables fundamentales. Dado que entre los modelos hay diferencias también entre el número de ítems (variables medidas) y el número de factores (variables latentes) no parece legítimo establecer todas las comparaciones posibles, sino especialmente las mostradas en la tabla 28 (modelos con 30 ítems, idéntico modelo de medida de las variables medidas e idéntico número y definición de variables latentes).

Los resultados obtenidos parecen indicarnos que los 3 modelos (EM-1.2.; EM-2 y EM-3) permiten aceptables explicaciones de las relaciones estructurales establecidas. **El modelo 3**, más cercano a un modelo predictivo, con 3 variables exógenas correlacionadas y sólo 2 variables endógenas, parece menos aceptable teóricamente, por ser menos parsimonioso. En todo caso, es interesante observar en él la **plausibilidad del efecto mediacional de la variable latente «Evaluación-Exámenes»**.

De los otros dos modelos (EM-1.2. y EM-2), parece mejor el primero, algo menos parsimonioso, pero **con mejores valores de ajuste** (similares indicadores de Bondad de ajuste en ambos modelos, pero mejor coeficiente de Determinación Total para el modelo EM-1) **lo que permite proponerlo como mejor modelo**.

En todo caso, no deben descartarse definitivamente los otros dos modelos, especialmente el **modelo EM-1.2.**, más parsimonioso (con 1 sola variable exógena y 4 endógenas) y que **presenta un plausible valor mediacional de la variable «Evaluación-Exámenes»**. Habrá que realizar estudios diferenciales en cada submuestra (en cada carrera, en cada tipo de materia, etc...) al objeto de comprobar que se mantienen estos resultados o por el contrario cambian en la línea de que cada modelo sea superior o inferior a otro en función de la submuestra utilizada.

3. REFERENCIAS BIBLIOGRÁFICAS

Abalde, E; De Salvador, X; González, R. y Muñoz, J.M. (1995). «Análisis de la evaluación de la docencia universitaria por los alumnos en la universidad de la Coruña (1993-94)». VII Seminario Nacional de Modelos de Investigación Educativa. Valencia.

- Abbot, R.D.; Wulff, D.H.; Nyquist, J.D.; Ropp, V.A. y Hess, C.W. (1990). «*Satisfaction with processes of Collecting Student Opinions about instruction: the student Perspective*». *Journal of Educational Psychology*, vol. 82 (2) (pp. 201–206).
- Abrami, P.C.; Cohen, P.A. y D'apollonia, S. (1990). «*Validity of Student Ratings of Instruction: What we know and what do not*». *Journal of Educational Psychology*. 82 (2) (pp. 219–231).
- Acherman, H. (1992). «*Evaluación de la calidad por el profesorado*». Actas del Congreso Internacional de Universidades. Universidad Complutense de Madrid. Madrid.
- Apodaka, P. y Otros (1990). «*Experiencia evaluativa en la UPV/EHU: consideraciones en torno a la puesta en marcha de un proyecto*». *Revista Española de Pedagogía* n° 186 (pp. 327–336).
- Arbizu, F. (1994). *La función docente del profesor universitario*. Universidad del País Vasco. Tesis Doctoral. Bizcaia.
- Asensio, I.I. (1993). *La medida del clima en instituciones de educación superior*. Tesis Doctoral. Edit. Universidad Complutense. Madrid.
- Baxter, E.P. (1991). «*The TEVAL Experience, 1983–1988: the impact of student evaluation of teaching scheme on university teachers*». *Studies in Higher Education*, n° 16 (pp. 151–178).
- Benedito, V.; Cabrera, F; Hernández, F; Mercade, F y Rodríguez Espinar, S. (1989). «*La evaluación del profesor universitario*». *Revista de Educación* n° 290 (pp. 279–291).
- Bollen, K.A. y Scott-Long, J. (Eds.). *Testing structural equation models*. Sage Pb. Newbury Park.
- Borrego, A. (1992). «*Evaluación institucional universitaria: el nivel académico*». Actas del Congreso Internacional de Universidades. Universidad Complutense. Madrid.
- Cajide, J. (1994). «*Análisis factorial confirmatorio de las características de la calidad docente universitaria*». *Bordón* 46 (4) (pp. 389–405).
- Castejón, J.L.; Carda, R.H. y Vera, M^a. Y. (1991). *Enseñanza universitaria: diseño y evaluación. Cuestiones teóricas y estudio aproximativo*. Universidad de Alicante.
- Centra, J.A. (1974). «*The relationship between student and alumni rating of teachers*». *Educational and Psychological Measurement* 34 (pp. 321–326).
- Centra, J.A. (1977). «*Student ratings of instruction and the relationship to student learning*». *American Educational Research Journal*. 14 (pp. 17–24).
- Centra, J.A. (1980). *Determining faculty performance*. Jossey-Bass. San Francisco.
- Cohen, P.A. (1981). «*Student ratings of instruction and student achievement: a meta-analysis of multisection validity studies*. *Review of Educational Research* 51 (pp. 281–309).
- Cowan, J. (1985). «*Effectiveness and efficiency in higher education*». *Higher Education* 14 (pp. 235–239).
- Dalhlöf, W. (1990a). *Evaluation de l'enseignement: pratiques et development. Vers nouveau modele*. OCDE: Centre pour la recherche et l'innovation dans Paris.
- Dalhlöf, W. (1990b). «*Practice and evidence in the evaluation of teaching*». Ed. por Dalhlöf et al., en Report of the IMHE study group on Evaluation in Higher Education. OCDE. Paris (June 1990) (cop. 5; pp. 139–161).
- Dalhlöf, W. (1990c). «*Towards a new model for the evaluation of teaching: an interactive process-centred approach*». Ed. por Dalhlöf et al, en Report of the IMHE study group on Evaluation in Higher Education. OCDE. Paris (June, 1990), (cap. 6, pp. 163–215).

- Dunkin, M. y Barnes, J. (1986). «*Research on higher education*». En Wittrock, P.: *Handbook of research on teaching*. McMillan. New York. (pp. 754–777).
- Entwistle, N. J. y Ramsden (1983). *Understanding student learning*. Croom Helm. London.
- Entwistle, N.J. y Tait, H. (1990). *Approaches to learning, evaluations of teaching, preferences for contrasting academic environments*. Higher Education. N° 19 (pp. 169–194).
- Escudero Escorza, T. (1989). «*Aproximación pragmática a la evaluación de la universidad*». *Revista de Investigación Educativa*. Vol. 7 (13) (pp. 93–112).
- Escudero Escorza, T. (1993). «*Enfoques modélicos en la evaluación de la enseñanza universitaria*». III jornadas nacionales de Didáctica Universitaria. Universidad de Gran Canaria (pp. 3–59).
- Fernández Díaz, M.J. (1986). *Investigación evaluativa de instituciones universitarias*. Tesis Doctoral. Universidad Complutense. Madrid.
- Fernández, J. (1995). «*Metaevaluación del profesorado de la calidad docente*». *Gaceta Complutense*. Oct. 95 (n° 111) (pp. 16–19).
- Fernández, M.R.; Jornet, J.M.; Pérez Carbonell, A. y González Such, J. (1995). «*Estudio del cuestionario de evaluación de la docencia a partir de las opiniones de los estudiantes: I y II cuestionario planteado; estudio metodológico y estudio factorial*». *Comunicaciones al VII seminario de Modelos de Investigación Educativa*. Valencia (pp. 300 a 308).
- Franke-Wikberg, S. (1990). «*Evaluating education quality on the instructional level*». *Higher Education Management*, 2 (3) (pp. 271–292).
- Frey, P.W.; Leonard, D.W. y Beatty, W.N. (1975). «*Student ratings of instruction: validation research*». *American Educational Research Journal* 12 (pp. 327–336).
- García Garrido, J.L. (1992). «*El profesorado universitario: un análisis comparativo de los modelos de formación*». *Actas del Congreso Internacional de Universidades*. Universidad Complutense. Madrid.
- García Ramos, J.M. (1986). «*Validez de constructo en el ámbito pedagógico*». *Revista Española de Pedagogía*. N° 174 (pp. 535–554).
- García Ramos, J.M. (1987). *El análisis factorial confirmatorio aplicado a la investigación pedagógica no-experimental*. *Bordón*, 267 (pp. 245–267).
- García Ramos, J.M. (1988). *Modelos exploratorios y confirmatorios en la investigación pedagógica no experimental*. *Rev. CC. de la Educación* n° 136. Oct–Dic. (pp. 423–450).
- García Ramos, J.M. (1989). «*Evaluación de la eficacia docente*» En Pérez Juste y García Ramos: *Diagnóstico, Evaluación y Toma de Decisiones* (Cap. 22). Rialp. Madrid.
- García Ramos, J.M. (1991). «*Recursos metodológicos en la evaluación de programas*». *Bordón* 43 (4). (pp. 461–476).
- García Ramos, J.M.; Sánchez, F.; Albert, C. y Salord, M. (1993). *Proyecto Educativo del Centro Universitario Francisco de Vitoria*. Publicación del Centro Universitario Francisco de Vitoria. Madrid.
- García Ramos, J.M. (1995). *La Evaluación Institucional a través del departamento de Recursos Humanos de una Institución Universitaria Privada*. *Bordón* 47(1), Enero–Feb. 1995 (pp. 17–30).
- García Ramos, J.M. y Congosto, E. (1995). «*Un sistema de evaluación institucional en la universidad*». *Comunicación al VII Seminario Nacional del Modelos de Investigación Educativa*. Valencia (pp. 264–267).

- García Ramos, J.M.; Fernández Díaz, M.J. y Congosto Luna, E. (1995). «*La evaluación de un modelo de tutoría académica en la universidad*». Comunicación al VII Seminario de Modelos de Investigación Educativa. Valencia. (pp. 268–271).
- García Ramos, J.M. (1997a). *Valoración de la Competencia Docente del Profesor Universitario. Una Aproximación Empírica*. Revista Complutense de Educación. Vol. 7 (3). Madrid.
- García Ramos, J.M. (1997b). *Análisis factorial confirmatorio a la Valoración del Constructo Competencia Docente del Profesor Universitario*. Bordón. Vol. 49 (4) (pp. 361–391). Madrid.
- Helmántica, (GRUPO de la Universidad de Salamanca) (1995). «*Elaboración de un instrumento de evaluación de las actitudes de los universitarios hacia la universidad, los alumnos y los profesores*». Comunicación al VII Seminario Nacional de Modelos de Investigación Educativa. Valencia. (pp. 313–318).
- Howard, R. D.; Nichols, J.O. y Gracie, L.W. (1987). «*Institutional research support of the self study*». En Muffo, J.A. y McLaughlin, G.W. (1987. A primer on Institutional research. Air. Florida Univ. Florida.
- Jordan, T.E. (1989). *Measurement and evaluation in higher education*. Falmer Press. London.
- Jöreskog, K.G. (1974). «*Analyzing Psychological Data by structural analysis of Covariance Matrices*». En D.H. Krahts, R. C. y otros (ed.): *Contemporary Developments in Mathematical Psychology*. (Vol. 2).
- Jöreskog, K.G. (1977). «*Structural equation model in the social sciences: specification, estimation and testing*». En P.R. Krishanaiah (ed.). *Applications of statistics*. North–Holland, Amsterdam.
- Jöreskog, K.G. y Sorböm, D. (1976). COFAMM. Confirmatory factor analysis with model modification. User's Guide, Nacional Educational Resources. Chicago.
- Jöreskog, K.G. y Sorböm, D. (1981). LISREL V: Analysis of linear structural relationship by method of maximum likelihood: user's Guide. Internacional Educational Services, Chicago.
- Jöreskog, K.G. y Sorböm, D. (1983). LISREL VI: Analysis of linear structural relationship by maximum likelihood and least square methods. International Educational services, Chicago.
- Jöreskog, K.G. y Sorböm, D. (1988). PRELIS: A program for multivariate data screening and data summarization: A preprocessor for LISREL. Univ. of Uppsala, Suecia (2ª Ed.).
- Jöreskog, K.G. y Sorböm, D. (1989). LISREL 7: A Guide to the Program and applications. Jöreskog–Sorböm (2ª Ed.).
- Jöreskog, K.G. y Sorböm, D. (1991). LISREL 7.20. A guide to the program and applications. Jöreskog–Sorböm–SPSS.
- Jornet, J.M. y Suárez (1988). *Evaluación de la docencia en Facultades de la Universidad de Valencia, a partir de las opiniones de los estudiantes*. Informes de investigación 3, 4, 5 y 6. Universidad de Valencia. Valencia.
- Jornet, J.M. y Suárez, J.M. y González Such, J. (1989). «*Estudio de validación de un cuestionario de valoración de la docencia universitaria para estudiantes en un conjunto homogéneo de centros de la universidad de Valencia*». Revista de Investigación Educativa, 7 (13) (pp. 57–92).

- Jornet, J.M.; González Such, J.; Suárez, J.M. y Pérez Carbonell, A. (1995) «Análisis de la consistencia de cuestionarios de opinión de estudiantes para evaluaciones docentes». Comunicación en VII Seminario Nacional de Modelos de Investigación Educativa. Valencia. (pp. 260–263).
- Jornet, J.M. (1993). «Enfoques de la evaluación universitaria». III Jornadas Nacionales de Didáctica Universitaria. Universidad de Gran Canaria (pp. 63–83).
- L'hommedieu, R.; Menges, R.J. y Brinko, K.T. (1988). *The effects of student ratings feedback to college teachers: a meta-analysis and review of research*. Center for the teaching professions. Northwestern University. Evanston.
- L'hommedieu, R.; Menges, R.J. y Brinko, K.T. (1990). «Methodological explanations for the modest effects of feedback from student ratings». *Journal of Educational Psychology*, vol. 83 (2) (pp. 232–241).
- Marlin, J.W. (1987). «Student perception of end-of-course evaluation». En *Journal of Higher Education* n° 58, 6.
- Marsh, H.W. y Overall, J.V. (1981). «The relative influence of course level, course type, and instructor on students evaluations of college teaching». *American Educational Research Journal*, 18 (pp. 103–112).
- Marsh, H.W. y Hocevar, D.W. (1983). «Confirmatory factor analysis of Multitrait–Multimethod matrices». *Journal of Educational Measurement* 20 (231–248).
- Marsh, H.W. (1982). «SEQ: a reliable, valid an useful instrument for collecting students evaluation of university teaching». *British Journal of Educational Psychology*. 52. (pp. 485–497).
- Marsh, H.W. (1983). «Multidimensional ratings of teaching effectiveness by student from different academic setting and their relation to student / course / instructor characteristics». *Journal of Educational Psychology* 75 (pp. 750–766).
- Marsh, H.W. (1984). «Students evaluations of university teaching: dimensionality, reliability, validity, potential biases and utility». *Journal of Educational psychology* 76. (pp. 707–754).
- Marsh, H.W. (1987). «Students evaluations of university teaching». Research findings methodological issues, and directions for future research. *Internacional Journal of Educational Research*, n° 11.
- Marsh, H.W. (1993). «The use of students evaluations and a individually structure intervention to enhance university teaching effectiveness». *Educational Research Journal*, 30 (1) (pp. 217–251).
- Martínez Arias, R. (1995). *Psicometría: Teoría de los tests psicométricos y educativos*. Síntesis. Madrid.
- Mateo, J. (1989). «La evaluación del profesorado universitario. Algunas reflexiones en torno al estado de la cuestión». En *informes de investigación evaluativa* n° 1. Consideraciones metodológicas sobre la evaluación y mejora de la docencia universitaria. Universidad de Valencia. (pp. 13–30).
- Mateo, M.A. y Fernández, J. (1992). «Análisis confirmatorio de la estructura dimensional de un cuestionario para la evaluación de la calidad de la enseñanza». *Rev. de Investigaciones Psicológicas* n° 11. Ed. Complutense. (pp. 73–82).

- McKeachie, W.J. (1990) «*Research on College Teaching: the historical background*». Journal of Educational Psychology. Vol. 80 (2). (pp. 189–200).
- Miguel, M. de (1988). «*Modelos de investigación sobre organizaciones educativas*». IV Seminario de Modelos de Investigación Educativa. Santiago de Compostela. Publicado en R.I.E. (1989). (pp. 21–56).
- Miguel, M. de (1990). «*Indicadores de calidad en la docencia universitaria*». Congreso sobre calidad de la Educación Universitaria. Puerto de Santa María. Cádiz.
- Miguel, M. de (1991). «*Utilización de indicadores en la evaluación docente universitaria*». En Miguel, M. de; Mora, J.G. y Rodríguez Espinar, S. (1991). *La evaluación de las Instituciones Universitarias*. Secretaría General del Consejo de Universidades. Madrid.
- Miller, R.I. (1987): *Evaluating faculty for promotion and tenure*. Jossey Bass. San Francisco.
- Millman, J. (ed.) (1981). *Handbook of teacher evaluation*. Sage. Beverly Hills.
- Mora, J.G. (1991). *Calidad y rendimiento en las instituciones universitarias*. Consejo de Universidades. Madrid.
- Mulaik, S.A. y Otros (1989). «*An evaluation of goodness of fit indices for structural equation models*». Psychological Bulletin, 105 (pp. 430–445).
- Murray, H.G. (1984). «*The impact of formative and summative evaluation of teaching in North American Universities*». Assessment and Evaluation in Higher Education, nº 9. (pp. 117–132).
- Neave, G. (1992). «*Hacia el gusto por la calidad: evaluación y universidad de masas*». Congreso Internacional de Universidades. Universidad Complutense. Madrid.
- Orden, A. de la (1985). «*Modelos de evaluación universitaria*». Revista Española de Pedagogía. (pp. 169–170 y 521–537).
- Orden, A. de la (1988). «*Calidad de los Centros Educativos. Asunto para un Congreso*». Borden 40 (2).
- Orden, A. de la (1992). «*Calidad y Evaluación de la Enseñanza Universitaria*». Actas del Congreso Internacional de Universidades. Universidad Complutense. Madrid.
- Oso, J.M. (1995). *Los indicadores de rendimiento en la evaluación institucional universitaria*. ICE. Universidad de Zaragoza. Zaragoza.
- Pérez Juste, R. (1986). «*Análisis y valoración del modelo español de educación superior a distancia*». Revista de Investigación Educativa. Vol. 4 (7).
- Pérez Juste, R. y García Ramos, J.M. (1989). *Diagnóstico, Evaluación y Toma de decisiones*. Rialp. Madrid.
- Ramsden, P. (1991). «*A performance indicator of teaching quality in higher education*». The experience questionnaire. Studies in Higher Education, nº 16 (pp. 129–150).
- Ramsden, P. y Entwistle, N.J. (1981). «*Effects of academic departament on students approaches to studying*». Britnish Journal of Educational Psychology 51 (pp. 368–383).
- Rodríguez Diéguez, J.L. (1992). «*Evaluación y formación del profesorado universitario: el caso español*». Actas del Congreso Internacional de Universidades. Universidad Complutense. Madrid.
- Rodríguez Espinar, S. (1986). «*Predicción y explicación del rendimiento relativo en función de los factores de personalidad del H.S.P.Q.*». Investigación presentada al Concurso para la Cátedra de la Universidad de Barcelona. Barcelona.

- Rodríguez Espinar, S. (1989). «*La evaluación del profesorado universitario. Análisis de la opinión del profesorado de la universidad de Barcelona*». En informes de Investigación Evaluativa nº 1. Consideraciones metodológicas sobre la evaluación y mejora de la docencia universitaria. Universidad de Valencia. (pp. 211–235).
- Rodríguez Espinar, S. (1990). «*Evaluación e innovación universitaria. ¿Por qué y para qué?*». En I. Coloquio Internacional: La Pedagogía Universitaria. Un reto en la Enseñanza Superior. Universidad de Barcelona. (Oct. 1990). Barcelona.
- Rodríguez Espinar, S. (1991a). «*Dimensiones de la calidad universitaria*». Ponencia presentada en el I Congreso Internacional sobre la calidad de la educación universitaria. 6–8 de marzo. Puerto de Santa María. Cádiz.
- Rodríguez Espinar, S. (1991b). «*Calidad universitaria: un enfoque institucional y multidimensional*». En, Miguel, M. de; Mora, J.G. y Rodríguez Espinar, S. (1991). *La Evaluación de las Instituciones Universitarias*. Secretaría General del Consejo de Universidades. Madrid.
- Rodríguez Espinar, S. (1993). «*Experiencias españolas de evaluación de la enseñanza universitaria y nuevas perspectivas. Otro punto de vista*». III Jornadas Nacionales de Didáctica Universitaria. Universidad de Las Palmas de Gran Canaria. (pp. 111–132).
- Rodríguez Espinar, S. (1996): «*Evaluación institucional y planificación universitaria*». Ponencia en I. Congreso Internacional de Evaluación de la Calidad de los Servicios en la Empresa. (Febrero 1996). Madrid.
- Salvador, L. y Sanz, J.J. (1988). *Evaluación de la docencia 1987–88. Opiniones del Alumno*. Vicerrectorado de Ordenación Académica. Universidad de Cantabria.
- Sánchez Carrión, J. J. (1984). *Introducción a las técnicas de análisis multivariante aplicadas a las CC. Sociales*. C.I.S. Madrid.
- Scott Long, J. (1987). *Confirmatory factor analysis*. Sage Univ. Paper. Nº 33. Beverly Hills. California.
- Scott Long, J. (1987). *Covariance Structure Models*. Sage Univ. Paper. Nº 34. Beverly Hills. California.
- Tanaka, J. S. (1993). «*Multifaceted conceptions of fit in structural equation models*». En Bullen, K.A. y Scott-Long, J. (ED.). *Testing structural equation models*. Sage Pb. News Bury Park.
- Tejedor, F.J. (1985). «*Problemática de la enseñanza universitaria*». *Revista de Investigación Educativa*. Vol. 3 (6) (pp. 322–337).
- Tejedor, F.J., Castro, C. y Mínguez, C. (1988). «*Evaluación del profesorado universitario por los alumnos*». *Studia Pedagógica*. nº 20 (pp. 73–134).
- Tejedor, F.J. (1990). «*La evaluación del profesorado en la universidad de Santiago*». *Revista Española de Pedagogía*, nº 186 (pp. 337–362).
- Tejedor, F.J. y Montero, L. (1990). «*Indicadores de la calidad docente para la evaluación del profesor universitario*». *Revista Española de Pedagogía*, nº 186. (pp. 260–279).
- Tejedor, F.J. (1993). «*Experiencias españolas de evaluación de la enseñanza universitaria y nuevas perspectivas*». III Jornadas Nacionales de Didáctica Universitaria. Universidad de las Palmas de Gran Canaria. (pp. 85–109).
- Tourón, J. (1989). *La validación de constructo: su aplicación al CEED, (cuestionario de evaluación de la eficacia docente)*. *Bordón* 1989 (pp. 735–756).

- Westerheijden, D.F. (1991). «*La evaluación de la universidad y su contexto político: gestión de la calidad y toma de decisiones en la educación superior*». En. Miguel, M. de; Mora, J.G. y Rodríguez, S. (1991). *Evaluación de las Instituciones Universitarias*. Secretaría General del Consejo de Universidades. Madrid.
- Whitely, S. y Dayle, K.O. (1976). «*Implicit theories in student rating*». *American Educational Research Journal*, 13 (pp. 241–253).
- Villa, A. y Morales, P. (1993). *La evaluación del profesor. Una visión de los principales problemas y enfoques en diversos contextos*. Nº 14 de Estudios y Documentos. Departamento de Educación, Universidades e Investigación. Gobierno Vasco. Vitoria.
- Wolf, R.L. (1992). «*La realidad actual de la universidad: un análisis de la calidad de la educación universitaria*». *Actas del Congreso Internacional de Universidades*. Universidad Complutense. Madrid.

TRABAJOS METODOLÓGICOS

COMPARACIÓN DE MÉTODOS ESTADÍSTICOS PARA LA EVALUACIÓN DEL FUNCIONAMIENTO DIFERENCIAL DE LOS ÍTEMS

Javier Gil Flores*, Eduardo García Jiménez* y Gregorio Rodríguez Gómez**

* Área MIDE de la Universidad de Sevilla

** Área MIDE de la Universidad de Cádiz

RESUMEN

En este trabajo llevamos a cabo un estudio comparativo de cuatro métodos para la detección del funcionamiento diferencial de los ítems: el método delta, uno de los más usados en el contexto de la teoría clásica, los métodos de las diferencias de proporciones y de chi cuadrado sumado, ambos basados en tablas de contingencia, y el contraste de parámetros b obtenidos al estimar la curva característica del ítem. Los datos utilizados corresponden a una prueba de comprensión de textos administrada a 519 alumnos de Educación Secundaria, y la variable diferenciadora de los grupos ha sido el sexo. Tras la aplicación de las diferentes técnicas, obtenemos una cierta convergencia entre los resultados alcanzados, dato que iría en contra de la hipótesis superioridad de unos métodos sobre otros.

ABSTRACT

A comparative study over four methods used to detect the differential item functioning is the focus of this paper. These methods are: the delta method (frequently used in the classic theory's context), the methods of proportion's differences and the chi square (both based on the contingency's table), and the contrast of the b parameters, obtained when estimating the characteristic curve of the item. The data used has been accomplished from a test about text's comprehension administered to 519 Secondary Education's students, being sex the differencing variable of the groups. After the application of these different techniques, the results obtained are similar.

INTRODUCCIÓN

Un tema que ha llegado a constituirse en foco de interés para investigadores y constructores de tests es el relativo al sesgo de los instrumentos de medida, tal y como refleja la profusión de trabajos que en los últimos años se han centrado en este tema. De acuerdo con el concepto habitualmente manejado, un instrumento de medida está sesgado cuando no ofrece la misma medida para dos sujetos o grupos de sujetos que cuentan con un nivel similar en el atributo medido, sino que sistemáticamente perjudica a alguno de ellos. Este mismo concepto podría trasladarse a cada uno de los ítems que componen el instrumento. Es decir, un ítem está sesgado cuando sujetos con la misma competencia y pertenecientes a distintas subpoblaciones no cuentan con el mismo grado de acierto al responder el ítem. El problema del sesgo representa un aspecto clave, no sólo desde el punto de vista social, dada la injusticia que representa para los colectivos que se ven perjudicados, sino también desde la óptica psicométrica, por el impacto negativo de los ítems sesgados sobre la validez de la prueba.

Ahora bien, no debe confundirse el sesgo del ítem con diferencias reales en el rendimiento de los grupos. El hecho de que hombres y mujeres, por ejemplo, tengan distinta tasa de aciertos en un ítem no significa que el ítem esté sesgado; puede ocurrir que el ítem discrimine entre ambos grupos por existir entre ellos una diferencia real en la capacidad medida. Al tratar de determinar el sesgo, lo que se persigue será precisamente separar las diferencias reales de las que son generadas por el propio instrumento de medida.

Generalmente, el estudio del sesgo de los ítems se ha vinculado a diferenciaciones de los sujetos desde el punto de vista social, cultural, racial, sexual, religioso, económico, geográfico: negros-blancos, mujeres-hombres, ricos-pobres,... Normalmente se alude a dos grupos, a los que se denomina grupo focal y grupo de referencia, que son respectivamente el grupo minoritario perjudicado o favorecido por el ítem y el grupo de comparación, respecto al cual se produce la diferencia observada en el grupo focal. También sería posible considerar el sesgo comparando grupos de similar tamaño o tomando más de dos grupos.

A veces, el sesgo se debe a la presencia de una variable que contamina la respuesta del individuo. Por ejemplo, si en un test de inteligencia incluimos una pregunta en la que se emplea un lenguaje complejo y elevado, ese elemento estará midiendo no sólo la inteligencia sino la capacidad de comprensión lectora de los sujetos, y resultará sesgado contra los sujetos poco competentes en comprensión lectora. Existen otras múltiples causas de sesgo de los ítems; en realidad, es difícil encontrar pruebas que no estén sesgadas.

La mejor forma de prevenir el sesgo en los ítems consiste en llevar a cabo un cuidadoso análisis del contenido de los mismos por parte de expertos. Sin embargo, cabe la posibilidad de realizar algunos análisis estadísticos mediante los cuales podemos detectar ítems que escaparon al examen previo. Los métodos para evaluar el sesgo se han basado en la comparación de la diferencia de respuesta en los dos grupos registrada para el ítem y la diferencia de respuesta que permanece constante a lo largo del test. Si se asume la unidimensionalidad del test, todos sus ítems miden el mismo cons-

tructo, y por tanto la diferencia en el grado de acierto en dos grupos distintos habría de mantenerse constante a lo largo de todos ellos. Si en un ítem se registran diferencias de acierto que se apartan de las que se dan en la globalidad de la prueba, podremos sospechar que se trata de un ítem con sesgo. Sin embargo, cabría la posibilidad de que todos los ítems contaran con un mismo sesgo y ese ítem destacara sobre los demás. Es decir, se podrían detectar ítems que resultan especialmente sesgados respecto al conjunto de ítems. Por este motivo resulta preferible, desde el punto de vista metodológico, hablar de técnicas para la evaluación del funcionamiento diferencial de los ítems. En la literatura psicométrica se suele recoger este concepto mediante la sigla DIF, que proviene de la expresión inglesa *Differential Item Functioning*.

OBJETIVO DEL ESTUDIO

Existen múltiples métodos estadísticos para evaluar el DIF. Los estudios comparativos y las simulaciones dirigidos a determinar la eficacia de diferentes métodos han sido frecuentes en los últimos años (Cohen y Kim, 1993; Zwick y otros, 1994a, 1994b), y son innumerables las variantes propuestas o las adaptaciones a casos concretos, formuladas a partir de los métodos usados habitualmente (Diamond, 1992; Miller y Spray, 1993; Oshima y otros, 1994; Mazor y otros, 1994, 1995; Nandakumar, 1994; Pang y otros, 1994; Hanson y Feinstein, 1995,...).

A pesar de tales estudios, no existe acuerdo sobre qué métodos resultan más adecuados. Para algunos, los métodos clásicos se ven superados por las nuevas técnicas basadas en la teoría de respuesta a los ítems (TRI), mientras que otros ven en este tipo de procedimientos problemas conceptuales y exigencias teóricas que no siempre se cumplen en la práctica. En realidad, ningún método resulta suficiente por sí mismo. La primera de las directrices ofrecidas por Hambleton y otros (1993) de cara a llevar a cabo estudios sobre el DIF apunta en este sentido, afirmando que no existe un único método capaz de garantizar la detección de todos los ítems de un test afectados por un funcionamiento diferencial.

Ante la diversidad de opiniones, en este trabajo hemos pretendido llevar a cabo una comparación de distintos métodos a fin de comprobar la convergencia o, por el contrario, la discrepancia entre los resultados a que nos conducen. Con este objetivo, aplicaremos diferentes técnicas para la evaluación del DIF a los ítems de una prueba de comprensión de textos, examinando el posible sesgo que éstos presentan en función de la variable sexo. Hemos tomado el factor diferenciador sexo por ser ésta una variable comúnmente considerada en los estudios sobre DIF y por las ventajas prácticas que ofrece, ya que la diferenciación por sexos suele dar lugar a dos subgrupos de examinados de tamaño parecido, evitando que el reducido número de sujetos en alguno de ellos impida la aplicación de algunos métodos de análisis, basados en la TRI, que exigen contar con un mínimo tamaño muestral de cara a la estimación de parámetros.

Por tanto, se tratará de identificar, utilizando diferentes métodos de análisis, los ítems de la prueba que pudieran favorecer o perjudicar sistemáticamente a las alumnas (grupo focal) frente a los alumnos (grupo de referencia) al medir su capacidad para

la comprensión de textos. La comparación de los resultados nos permitirá valorar la coincidencia o discrepancia entre los diferentes métodos usados.

DATOS UTILIZADOS

Los datos utilizados en este estudio han sido extraídos de una investigación que tenía por objetivo la construcción de un test adaptativo computerizado (TAC), cuyo fin era medir la capacidad de los alumnos de Educación Secundaria para comprender información escrita en el área de Geografía e Historia (García y otros, 1993). Aquí hemos tomado una de las 49 pruebas en que fue fragmentado el banco inicial de ítems, de cara a su calibración conforme a un modelo logístico de 3 parámetros. La calibración de un banco de ítems para la construcción de un TAC requiere la aplicación previa de cada una de las pruebas en el formato convencional no adaptativo de lápiz y papel. Los datos obtenidos con una de estas pruebas son los que utilizaremos en la comparación de métodos para la evaluación del DIF.

La prueba consistía en la presentación de diversos textos históricos o relativos a temas de geografía física, humana y económica, que daban paso a la formulación de una serie de cuestiones o ítems sobre su contenido. En cada ítem se ofrecían cuatro opciones de respuesta, de las cuales una sola era correcta, funcionando las tres restantes como distractores. Las puntuaciones correspondientes a cada ítem fueron dicotómicas, asignando 1 en caso de acierto y 0 en caso de error, y la suma de todas ellas se tomó como puntuación alcanzada en la prueba.

El número de examinados ascendió a 519 sujetos, de los cuales 219 eran alumnos y 300 alumnas. Todos ellos cursaban estudios de BUP o niveles equivalentes de la nueva Educación Secundaria, en el I.B. Caballero Bonald, de Jerez de la Frontera, y en el I.E.S. José Luis Tejada, de El Puerto de Santa María, ambos en la provincia de Cádiz. La administración tuvo lugar en el segundo trimestre del curso académico 1994/95. Una vez eliminados, de cara a mejorar la calidad psicométrica de la prueba, los ítems que presentaban una dificultad (o facilidad) extrema y no alcanzaban un nivel mínimo de discriminación (correlación biserial puntual inferior a 0.30), el instrumento quedó constituido por 20 ítems, los cuales han sido objeto del presente estudio.

MÉTODOS PARA LA EVALUACIÓN DEL DIF

La identificación de los ítems afectados por DIF se ha realizado recurriendo a un total de cuatro métodos estadísticos. Hemos seleccionado el método delta, por tratarse de uno de los más ampliamente usados en el contexto de la teoría clásica, y también los métodos de la diferencia de proporciones y de chi cuadrado sumado, ambos basados en el análisis de tablas de contingencia. Por último, hemos considerado una técnica para la identificación del DIF basada en los modelos de la TRI; concretamente, se trata de la comparación estadística de los parámetros b obtenidos al estimar la curva característica del ítem para el grupo focal y el grupo de referencia. Tras un breve comentario sobre cada uno de los métodos, presentaremos los resultados obtenidos.

Método delta

Uno de los procedimientos más usados en el contexto de la teoría clásica es el método delta (Angoff, 1982; Angoff y Ford, 1973). El principio básico consiste en encontrar los ítems con mayor discrepancia entre los índices de dificultad calculados para los dos grupos. El método delta supone transformar las proporciones de acierto (índice de dificultad) para los dos grupos en puntuaciones delta (puntuaciones típicas derivadas, con media 13 y desviación típica 4). Por tanto, en primer lugar habrá que transformar los índices de dificultad (proporciones de acierto p) en puntuaciones z correspondientes al percentil $1-p$ de la distribución normal y posteriormente trasladarla a la escala de las puntuaciones delta. Al obtener la puntuación z a partir de $1-p$ y no a partir de p , conseguimos que las puntuaciones altas correspondan a los ítems de mayor dificultad y las puntuaciones bajas a los de menor dificultad. Los índices de dificultad (p_1 y p_2), las puntuaciones z_{1-p} correspondientes a los mismos (z_1 y z_2) y las puntuaciones delta (Δ_1 y Δ_2) para los 20 ítems de la prueba aparecen recogidos en la tabla 1.

Tabla 1
RESULTADOS DE LA APLICACIÓN DEL MÉTODO DELTA

Ítem	p_1	p_2	z_1	z_2	Δ_1	Δ_2	d
1	0.75	0.66	-0.67	-0.41	10.32	11.36	-.24
2	0.50	0.34	0.00	0.41	13.00	14.64	-.74
3	0.15	0.22	1.04	0.77	17.16	16.08	.66
4	0.86	0.75	-1.08	-0.67	8.68	10.32	-.51
5	0.79	0.74	-0.81	-0.64	9.76	10.44	.01
6	0.79	0.71	-0.81	-0.55	9.76	10.80	-.21
7	0.24	0.21	0.71	0.81	15.84	16.24	-.15
8	0.52	0.43	-0.05	0.18	12.80	13.72	-.30
9	0.34	0.34	0.41	0.41	14.64	14.64	.15
10	0.76	0.76	-0.71	-0.71	10.16	10.16	.39
11	0.59	0.56	-0.23	-0.15	12.08	12.40	.10
12	0.38	0.34	0.31	0.41	14.24	14.64	-.07
13	0.92	0.92	-1.41	-1.41	7.36	7.36	.54
14	0.54	0.52	-0.10	-0.05	12.60	12.80	.14
15	0.48	0.34	0.05	0.41	13.20	14.64	-.63
16	0.36	0.34	0.36	0.41	14.44	14.64	.04
17	0.29	0.28	0.55	0.58	15.20	15.32	.05
18	0.37	0.35	0.33	0.39	14.32	14.56	.02
19	0.23	0.21	0.74	0.81	15.96	16.24	-.09
20	0.58	0.57	-0.20	-0.18	12.20	12.28	.23

Las puntuaciones delta de cada ítem son representadas gráficamente situando en abscisas la puntuación en uno de los grupos, y en ordenadas la puntuación alcanzada en el otro. De este modo, conseguimos un gráfico formado por tantos puntos como ítems posee el test (ver más adelante figura 1). Esta nube de dispersión tendrá la forma de una elipse; cuando los grupos provienen de una misma población, la elipse se estrechará aproximándose a una recta. Si los puntos se encuentran en línea recta, podemos afirmar que los ítems son insesgados, estando sesgados aquéllos que se apartan considerablemente de ella. No obstante, como afirma Muñiz (1992), el hecho de que los ítems se sitúen en línea recta no implica ausencia de sesgo. Podríamos pensar que una línea recta por debajo de la diagonal representa un sesgo contra el grupo representado en ordenadas, mientras que una recta por encima de la diagonal supondría un sesgo contra el grupo representado en abscisas. En realidad, el método delta evalúa la discrepancia entre sesgos, detectando aquellos ítems que se apartan en este rasgo de los restantes que constituyen la prueba.

Se han propuesto índices que permiten caracterizar el sesgo de los ítems sin necesidad de basarnos exclusivamente en la inspección visual del gráfico. Globalmente, el ajuste de los puntos a la recta puede valorarse a partir de la correlación de Pearson entre los dos grupos de puntuaciones delta. En este caso, el valor de la correlación asciende a 0.97, indicando un buen ajuste y por tanto un bajo DIF. Sin embargo, más que una consideración global sobre los 20 ítems de la prueba, nos interesa caracterizar a cada uno de ellos particularmente. Un modo de hacerlo se basa en calcular la distancia de cada punto al eje principal de la elipse, de forma que cuanto más grande sea esta distancia mayor será el sesgo del ítem en relación a los otros. Para calcularla, siguiendo a Angoff y Ford (1973), partimos del eje principal de la elipse, cuya ecuación corresponderá a la recta

$$\Delta_2 = a \cdot \Delta_1 + b$$

donde los valores de las constantes a y b son

$$a = \frac{S_{\Delta_1}^2 S_{\Delta_2}^2 + \sqrt{(S_{\Delta_2}^2 S_{\Delta_1}^2)^2 + 4 r_{\Delta_1 \Delta_2}^2 S_{\Delta_1}^2 S_{\Delta_2}^2}}{2 r_{\Delta_1 \Delta_2} S_{\Delta_1} S_{\Delta_2}}$$

$$b = \bar{\Delta}_2 - a \cdot \bar{\Delta}_1$$

El índice de distancia para un ítem j, cuyas puntuaciones delta en ambos grupos son D_{1j} y D_{2j} respectivamente, vendrá expresado por:

$$d = \frac{a \Delta_{1j} - \Delta_{2j} + b}{\sqrt{a^2 + 1}}$$

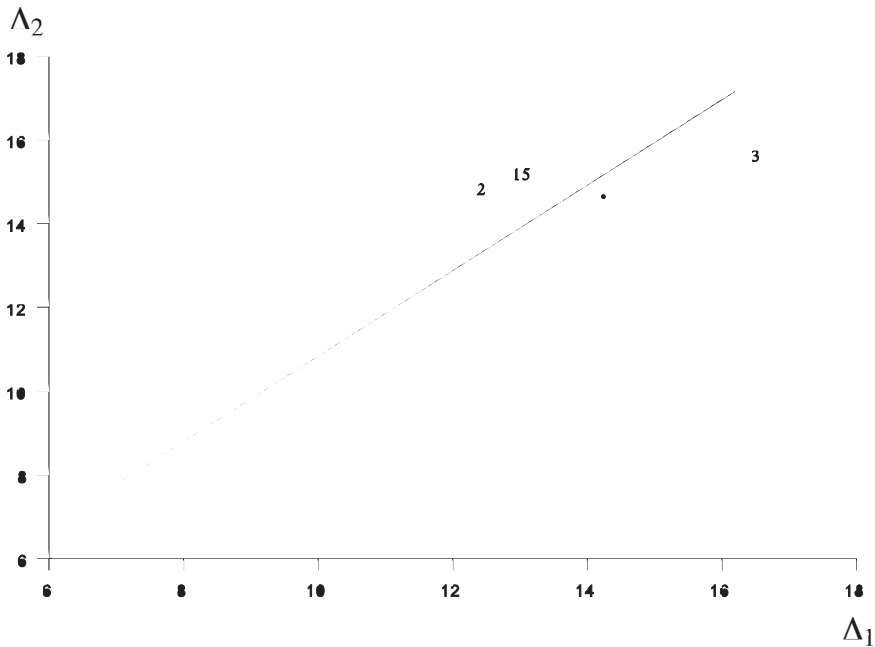
La ecuación para el eje mayor de la elipse que forman los puntos del diagrama de dispersión ha sido determinada en el caso de la prueba que nos ocupa, obteniendo como resultado:

$$\Delta_2 = 0.91 \cdot \Delta_1 + 1.57$$

En la figura 1, este eje ha sido trazado sobre el diagrama de dispersión para las puntuaciones delta, etiquetando por medio de su número a los ítems que más se distancian del eje.

Figura 1

DIAGRAMA DE DISPERSIÓN PARA LAS PUNTUACIONES DELTA, EJE MAYOR DE LA ELIPSE Y PUNTOS MÁS DISTANCIADOS DEL EJE



Calculando la distancia al eje para todos los ítems tenemos una visión global y podemos establecer, siguiendo criterios meramente descriptivos, puntos de corte a partir de los cuales descartar ítems que consideramos especialmente sesgados. Las distancias (d) quedaron recogidas en la tabla 1 mostrada anteriormente. Fijando en 0.60 el nivel de corte, los ítems que mayor DIF presentan son, en este orden, los números 2, 3 y 15. El signo positivo asociado a la distancia d indica que el ítem favorece al grupo focal, mientras que las distancias negativas corresponden a ítems que lo perjudican.

Una modificación propuesta por Camilli y Saphard (1994) al método de Angoff consiste en realizar una segunda estandarización de los índices de dificultad. En lugar de calcular la distancia de los puntos a la recta, tomamos las puntuaciones z_{1-p} y las estandarizamos respecto al grupo, a partir de la media y la desviación típica del conjunto de puntuaciones z_{1-p} . Es decir, la nueva puntuación z^* correspondiente a una puntuación z se obtendrá, por tanto, como:

$$z^* = \frac{z - \mu_z}{\sigma_z}$$

donde μ_z es la media de las puntuaciones z_{1-p} para todos los ítems de la prueba, y σ_z la desviación típica de las mismas. La diferencia $z_1^* - z_2^*$ entre las puntuaciones obtenidas para un mismo ítem en los grupos focal y de referencia puede ser tomada como medida del DIF. No obstante, el cálculo de este valor no altera el resultado obtenido basándonos en la distancia al eje; de nuevo, los valores más altos para la diferencia entre puntuaciones z estandarizadas corresponden a los ítems 2, 3 y 15, donde $z_1^* - z_2^*$ asciende a -0.48 , 0.50 y -0.41 respectivamente.

Método de la diferencia de proporciones

Este método, propuesto por Dorans y Kurlick (1983, 1986), se basa en el cálculo de la diferencia de proporciones en cada uno de los niveles en que previamente son subdivididos ambos grupos. Suponiendo que hemos dividido los grupos en k niveles de capacidad, podemos calcular la diferencia en la proporción de aciertos para cada uno de esos niveles. Si p_{2j} y p_{1j} son las proporciones de acierto en los grupos focal y de referencia para el nivel de capacidad j , la diferencia de proporciones vendrá dada por:

$$\Delta p_j = p_{2j} - p_{1j}$$

A partir de la suma ponderada de las diferencias de proporciones en los k niveles, se puede calcular el estadístico:

$$\text{DPE DIF} = \frac{\sum_{j=1}^k w_j \Delta p_j}{\sum_{j=1}^k w_j}$$

donde el coeficiente de ponderación w_j puede adoptar diferentes valores, según los propósitos de la investigación (Dorans y Holland, 1993). Se recomienda el uso de n_{2j} o

frecuencia en j de los sujetos del grupo focal, pues de esta manera se da más peso a las diferencias de proporciones registradas en niveles de capacidad donde los sujetos del grupo focal son más numerosos.

Los valores del estadístico DPE-DIF están comprendidos entre -1 y 1 . Siguiendo las pautas de interpretación ofrecidas por Dorans y Holland (1993), cuando encontramos valores que se sitúan, en valor absoluto, por debajo de 0.05 , podemos afirmar la inexistencia de un funcionamiento diferencial de los ítems; valores absolutos comprendidos entre 0.05 y 0.10 no son preocupantes, aunque aconsejan una inspección de los ítems; y valores por encima de 0.10 , en valor absoluto, hacen necesario un cuidadoso examen. Los valores positivos indicarán una ventaja para el grupo focal, mientras que los negativos reflejan una desventaja para este grupo.

La aplicación de este método a los ítems de la prueba analizada se ha basado en la división de los grupos en 6 niveles de capacidad. Todos los intervalos creados poseen una amplitud de 3 unidades. La frecuencia de sujetos en cada nivel de dificultad, tanto para el grupo focal como para el grupo de referencia, aparece recogida en la tabla 2.

Tabla 2
NÚMERO DE SUJETOS INCLUIDOS EN CADA NIVEL DE CAPACIDAD

Niveles	G. Referencia	G. Focal
1-3	3	5
4-6	29	48
7-9	63	110
10-12	63	82
13-15	43	37
16-18	18	18

Los valores de las proporciones de acierto en cada nivel, así como el estadístico DPE-DIF calculado para cada uno de los ítems, aparecen en la tabla 3. Considerando afectados de DIF aquellos ítems en los que el estadístico calculado se aproxima o supera en valor absoluto la cota de 0.10 , este método señala a los números 2, 3, 4 y 15, de entre los cuales únicamente el ítem 3 favorece al grupo focal, de acuerdo con el signo positivo asignado.

Método de chi cuadrado sumado

Uno de los métodos basado en el estadístico χ^2 , es el debido a Camilli (1979). Para cada ítem, se parte de la construcción de tablas referidas a cada nivel de capacidad, en la que se presentan las frecuencias de acierto y error para los grupos focal y de referencia (ver tabla 4).

Tabla 4
 TABLA DE CONTINGENCIA PARA EL ÍTEM i Y EL INTERVALO j

		Puntuación en el ítem i		
		Acierto (1)	Error (0)	
Grupo	Referencia	A_j	B_j	n_{Rj}
	Focal	C_j	D_j	n_{Fj}
		m_{1j}	m_{0j}	T_j

Los valores A_j , B_j , C_j y D_j representan las respectivas frecuencias en el intervalo j -ésimo al que se refiere la tabla. Los valores m_{1j} y n_{Rj} son frecuencias marginales, y el valor T_j se corresponde con la frecuencia total en el nivel de capacidad considerado. A partir de esta tabla es posible calcular el estadístico χ_j^2 , de acuerdo con la siguiente expresión:

$$\chi_j^2 = \frac{T_j}{n_{Rj} n_{Fj}} \left(\frac{A_j D_j}{m_{1j}} - \frac{B_j C_j}{m_{0j}} \right)^2$$

La suma de los valores χ_j^2 obtenidos para los k niveles de capacidad se distribuyen según k grados de libertad. Bastará comparar con el correspondiente valor crítico para decidir sobre la hipótesis nula de no existencia de DIF.

Utilizando los mismos intervalos de capacidad considerados al aplicar el método de la diferencia de proporciones, hemos calculado los correspondientes valores χ_j^2 , y a partir de éstos el valor de chi-cuadrado sumado de Camilli (χ_c^2). Todos ellos se recogen en la tabla 5. En el caso de ítems muy fáciles o muy difíciles no es posible calcular el valor de χ_j^2 en los intervalos de capacidad extremos, dado que existen celdas vacías o con frecuencias esperadas muy bajas. Los ítems en los que afirmamos la existencia de DIF serán aquéllos para los cuales el valor de χ_c^2 supera al valor crítico χ^2 , que para un nivel de significación $\alpha=0.05$ adopta los valores 5.99, 7.82 y 9.49 en los casos de 2, 3 y 4 grados de libertad respectivamente. Tales ítems han sido marcados en la tabla 5 mediante la colocación de un asterisco junto al valor de χ_c^2 .

Tabla 3
RESULTADOS DE LA APLICACIÓN DEL MÉTODO DE LA DIFERENCIA DE PROPORCIONES

Ítem	Nivel 1		Nivel 2		Nivel 3		Nivel 4		Nivel 5		Nivel 6		$\Sigma(n_i - \Delta_{ip}) / \Sigma n_i$
	P ₁	P ₂	P ₁	P ₂	P ₁	P ₂	P ₁	P ₂	P ₁	P ₂	P ₁	P ₂	
1	.000	.000	.551	.354	.587	.554	.825	.817	.930	.837	1.000	1.000	-.06
2	.333	.000	.137	.083	.317	.272	.476	.390	.860	.486	1.000	.888	-.11
3	.000	.000	.000	.062	.047	.109	.095	.158	.302	.513	.611	1.000	.10
4	.333	.600	.586	.645	.873	.681	.920	.743	.907	.945	1.000	1.000	-.10
5	.000	.200	.586	.625	.746	.736	.809	.719	.930	.891	.944	1.000	-.02
6	.333	.000	.586	.458	.682	.709	.873	.731	.930	.918	1.000	.888	-.06
7	.000	.000	.137	.020	.111	.163	.158	.195	.348	.405	.888	.722	.01
8	.000	.000	.103	.166	.460	.409	.619	.597	.697	.702	.722	.111	-.05
9	.333	.200	.241	.125	.269	.272	.365	.304	.372	.540	.611	1.000	.01
10	.000	.200	.551	.583	.634	.718	.825	.829	.930	.891	.944	.944	.04
11	.333	.400	.310	.208	.349	.500	.698	.670	.837	.648	.944	1.000	.01
12	.000	.000	.000	.062	.238	.190	.460	.451	.558	.648	.833	.944	.01
13	.333	.400	.827	.708	.873	.936	.968	1.000	.976	1.000	1.000	1.000	.02
14	.000	.200	.103	.250	.396	.390	.634	.646	.767	.756	.888	1.000	.03
15	.000	.000	.172	.104	.190	.209	.666	.390	.697	.621	.888	.944	-.09
16	.333	.000	.137	.270	.254	.236	.349	.353	.534	.621	.722	.666	.02
17	.000	.000	.069	.125	.190	.154	.222	.317	.465	.486	.833	.777	.02
18	.333	.200	.103	.208	.381	.254	.365	.402	.418	.405	.611	.944	.00
19	.333	.000	.103	.166	.095	.127	.254	.207	.279	.189	.611	.833	.01
20	.000	.200	.241	.187	.381	.381	.650	.743	.883	.973	.944	1.000	.03

Tabla 5

RESULTADOS DE LA APLICACIÓN DEL MÉTODO DE CHI CUADRADO SUMADO

Ítem	χ^2_2	χ^2_3	χ^2_4	χ^2_5	χ^2_c
1	2.881	.175	.016	.	3.07
2	.	.390	1.075	12.946	14.42 *
3	.	1.911	1.253	3.695	6.86
4	.273	7.840	7.561	.	15.67 *
5	.114	0.019	1.576	.	1.71
6	1.182	.134	4.335	.	5.65
7	.	.893	.320	.271	1.48
8	.590	.429	.068	.002	1.09
9	1.744	.001	.583	2.280	4.61
10	.073	1.293	.003	.	1.37
11	1.012	3.687	.126	3.771	8.60
12	.	.541	.011	.678	1.23
13	1.383	2.029	.	.	3.41
14	2.475	.005	.020	.012	2.51
15	.745	.086	10.893	.514	12.24 *
16	1.856	.067	.003	.612	2.54
17	.	.370	1.604	.036	2.01
18	1.417	3.044	.209	.014	4.69
19	.590	.402	.441	.887	2.32
20	.318	.000	1.480	.	1.80

Método del contraste de parámetros

Los métodos de estudio del sesgo basados en la TRI son muy populares entre los investigadores. La idea de partida es que si la probabilidad de acertar un ítem es similar para los sujetos del grupo focal y del grupo de referencia, entonces las respectivas curvas características del ítem en ambos grupos han de ser también similares. En consecuencia, no deben existir diferencias significativas entre las áreas delimitadas por las correspondientes CCI o entre los parámetros que describen a las curvas para los grupos focal y de referencia. Los métodos para comprobar la significación estadística de las diferencias entre parámetros son variados. Aquí llevaremos a cabo un contraste estadístico para la diferencia de parámetros b estimados en cada grupo, siguiendo el procedimiento que describimos a continuación.

Si denominamos b_F al parámetro b para el grupo focal y b_R al parámetro b para el grupo de referencia, someteremos a contraste la hipótesis nula de igualdad frente a la hipótesis alternativa que afirma la existencia de diferencias.

$$H_0: \Delta b = b_F - b_R = 0$$

$$H_1: \Delta b = b_F - b_R \neq 0$$

El estadístico de contraste utilizado (Lord, 1980) es el cociente entre la diferencia de parámetros y el error típico para las diferencias. Teniendo en cuenta que este error típico es

$$S_{\Delta b} = \sqrt{S_{b_F}^2 + S_{b_R}^2}$$

el estadístico de contraste será:

$$z_b = \frac{b_F - b_R}{\sqrt{S_{b_F}^2 + S_{b_R}^2}}$$

que se distribuye normalmente $N(0,1)$. La consulta de una tabla de valores para la distribución normal nos permitirá conocer la significación estadística en este contraste.

En esta prueba se requiere que los parámetros estimados en ambos grupos se encuentren en una misma métrica, antes de que la comparación se lleve a efecto. El procedimiento utilizado para obtener el valor z_b a partir de parámetros b_F y b_R expresados en una misma métrica ha sido el siguiente:

- a) Calibrar los ítems en el grupo de referencia, asumiendo un modelo logístico de 3 parámetros, y determinar los parámetros b y el error típico de los mismos. En la estimación de parámetros hemos adoptado el método de máxima verosimilitud marginal, debido a Bock y Lieberman (1970), e implementado mediante el programa XCALIBRE (Assessment Systems Corporation, 1995).
- b) Calibrar los ítems en el grupo focal, tomando los parámetros del grupo de referencia y sus errores típicos como definición de la métrica común. El método para la equiparación de parámetros ha sido el de la media y la desviación típica. Estudios basados en simulación han comprobado que, para muestras suficientemente amplias, el método de la media y la desviación típica y el método basado en la curva característica del test conducen a resultados similares en la posterior detección del DIF (Kim y Cohen, 1992).
- c) Calcular la medida DIF como diferencia entre los parámetros b de ambos grupos y valorar la significación de la diferencia.

La tabla 6 muestra los valores obtenidos para los parámetros b en los grupos focal y referencia, expresados en una métrica común, así como los errores típicos para los mismos. La última columna de esta tabla recoge el valor del estadístico de contraste Z_b . Como puede apreciarse, fijando un nivel de significación $\alpha=0.01$ (valores críticos $z_{\alpha/2}=\pm 2.57$) únicamente la diferencia correspondiente al ítem 2 resulta estadísticamente significativa. No obstante, si hubiéramos sido menos restrictivos ($\alpha=0.05$ y valores críticos $z_{\alpha/2}=\pm 1.96$) las diferencias de parámetros para ítems como el 3 y el 15, cuyo DIF fue detectado por métodos anteriores, también habrían resultado significativas. Los signos unidos a los valores del estadístico de contraste resultan coincidentes con los que adoptaban indicadores utilizados en otros métodos.

Tabla 6
RESULTADOS DE LA APLICACIÓN DEL MÉTODO DE LA DIFERENCIA DE PARÁMETROS

Ítem	b_F	b_R	S_{bR}	S_{bR}	Z_b
1	-0.59	-0.72	.141	.163	— .60
2	1.47	0.65	.222	.163	-2.98
3	1.68	2.58	.220	.380	2.05
4	-1.13	-1.56	.164	.186	-1.73
5	-1.15	-1.04	.173	.180	.44
6	-0.89	-1.04	.157	.167	— .65
7	2.29	2.00	.313	.266	— .71
8	1.53	0.67	.284	.204	-2.46
9	1.50	2.23	.222	.342	1.79
10	-1.24	-0.80	.159	.158	1.96
11	0.17	0.15	.172	.150	— .09
12	1.17	1.25	.184	.203	.29
13	-2.48	-2.04	.193	.203	1.57
14	0.27	0.39	.152	.160	.54
15	1.26	0.67	.195	.166	-2.30
16	1.75	1.71	.263	.254	— .11
17	1.96	1.88	.275	.260	— .21
18	1.58	1.90	.238	.289	.85
19	2.37	2.37	.321	.343	.00
20	-0.11	0.17	.123	.151	1.44

COMPARACIÓN DE RESULTADOS

Tras la aplicación de los diferentes métodos para la identificación de ítems que presentan DIF, vamos a calibrar el grado de acuerdo existente entre los resultados logrados con cada uno de ellos. La tabla 7 presenta una síntesis comparativa de las técnicas aplicadas, indicando que todos los ítems detectados lo son al menos por dos de las cuatro técnicas aplicadas. Además, los cuatro procedimientos de análisis han señalado el ítem 2 como ítem afectado de DIF y existe también amplio acuerdo en torno al ítem 15. A esto podemos sumar la coincidencia en el signo del sesgo, en los casos en que éste es considerado. Así, las tres técnicas que informan sobre el sentido del sesgo (métodos delta, de la diferencia de proporciones y del contraste de parámetros) convergen al indicar que los ítems 2, 4 y 15 perjudican a los examinados pertenecientes al grupo focal, mientras que el ítem 3 los favorece.

Los criterios establecidos para considerar un ítem sesgado pueden llegar a ser un tanto arbitrarios. En el caso del método delta, hemos fijado una distancia mínima de 0.60 al eje mayor de la elipse, aunque no existe ningún motivo que nos impidiera haber tomado un nivel inferior o superior como punto de corte. Los métodos de chi cuadrado sumado y del contraste de parámetros requieren que el analista fije un nivel de significación para decidir cuándo un ítem presenta DIF. También el método de la diferencia de proporciones deja un tanto abiertas las pautas o criterios de decisión. Por estas razones, más que etiquetar a los ítems como afectados o no afectados por un funcionamiento diferencial, consideramos adecuado hablar del grado en que presentan este rasgo.

Tabla 7

ÍTEMES QUE PRESENTAN FUNCIONAMIENTO DIFERENCIAL, DE ACUERDO CON LOS CUATRO MÉTODOS UTILIZADOS

Método	Ítem			
	2	3	4	15
Método delta (Angoff y Ford, 1973)	*	*		*
Diferencia de proporciones (Dorans y Kurlick, 1983,1986)	*	*	*	*
Chi-cuadrado sumado (Camilli, 1979)	*		*	*
Contraste de parámetros b (Lord, 1980)	*			

Cada uno de los métodos usados proporciona índices que pueden ser tomados como medidas del grado en que el ítem está afectado por DIF: la distancia al eje de la elipse (método delta), el estadístico DPE-DIF (método de la diferencia de proporciones), el valor medio que alcanza c^2_j en los distintos tramos de capacidad en que son divididos los grupos (método de chi cuadrado sumado) o el valor z_b (método del contraste de parámetros). En consecuencia, un estudio de la convergencia entre los distintos métodos podría apoyarse en la correlación que se da entre los respectivos indicadores de DIF, tomados todos ellos en valores absolutos. De este modo, podríamos obtener una medida de la convergencia entre métodos incluyendo en su cálculo a todos y cada uno de los 20 ítems que constituyen la prueba. La matriz de correlaciones intermétodos aparece en la tabla 8, reflejando una moderada o incluso alta convergencia entre los resultados a los que nos conduce cada método.

Tabla 8

CORRELACIONES ENTRE LOS VALORES DEL DIF ESTIMADOS POR DIFERENTES MÉTODOS

	Método delta	Diferencia de proporciones	Chi-cuadrado	Contraste de parámetros
Método delta	1.0000			
Diferencia de proporciones	.8455	1.0000		
Chi-cuadrado	.7129	.7661	1.0000	
Contraste de parámetros	.8561	.8561	.5147	1.0000

Los valores de la correlación de Pearson indican que las medidas de DIF obtenidas a partir del método delta son las que mayor covariación registran con las resultantes de los demás métodos, obteniéndose correlaciones de 0.8561, 0.8455 y 0.7129 con los métodos del contraste de parámetros, de la diferencia de proporciones y de chi cuadrado respectivamente. La correlación más elevada se ha registrado entre el método delta y el método de contraste de parámetros b . La coincidencia entre los resultados alcanzados por uno y otro método se confirma si tenemos en cuenta que el método delta destacaba a los ítems 2, 3 y 15 como los que registraban mayor DIF, y estos tres elementos se encuentran también entre los cuatro ítems que presentan una mayor diferencia de parámetros b .

CONCLUSIÓN

El desarrollo de los métodos para la evaluación del DIF ha surgido como respuesta a la necesidad de identificar ítems que presentan sesgo. No obstante, las técnicas ideadas cuentan con limitaciones que debemos tener presentes. En el caso del método delta, una crítica habitual es la posible confusión entre dificultad y discriminación. La dificultad diferencial no permanecerá constante a lo largo de los ítems de una prueba, a menos que todos presenten un mismo nivel de discriminación. Es posible que los ítems identificados por un funcionamiento diferencial se caractericen en realidad por un alto poder de discriminación. Por otra parte, el que todos los puntos representados en el gráfico se alineen supone ausencia de DIF, pero ello no implica ausencia de sesgo; cabría la posibilidad de que todos los ítems estén igualmente sesgados. Al aplicar este método tendríamos que partir del supuesto de que la mayoría de los ítems son carentes de sesgo. En general, los métodos basados en la teoría clásica, presentan un problema circular: para la valoración del sesgo es preciso partir de diferentes niveles de capacidad medidos por el propio test, pero si éste presenta sesgo, tales niveles no serían una referencia válida.

Los métodos basados en la TRI salvan este problema, al apoyarse en la comparación de las curvas características del ítem estimadas para cada subgrupo, en lugar de tomar índices de dificultad o proporciones de aciertos. Sin embargo, y a pesar de que se ha demostrado la utilidad práctica de estos métodos, existe en ellos una contradicción teórica: si se asume que la estimación de parámetros es invariante respecto de los sujetos, resultaría imposible que dos submuestras generasen dos curvas diferentes. Como afirman Muñoz y Hambleton (1992), la realidad es más boscosa que la teoría y lo cierto es que, a pesar de unas bases teóricas no todo lo sólidas que se desearía, estos métodos detectan con buena precisión los ítems sesgados.

Tal vez la discusión sobre la superioridad de unos métodos u otros pueda resultar estéril a la luz de resultados como los obtenidos en este trabajo. El alto grado de coincidencia en los valores a los que nos conducen los métodos para la evaluación del DIF bien nos permitiría su utilización indistinta. En particular, en el caso aquí estudiado, la convergencia entre las medidas DIF obtenidas mediante el método delta (el más usado en el contexto clásico) y el contraste de parámetros b ponen de manifiesto que los planteamientos apoyados en la teoría clásica y en la TRI no son, en la práctica, tan alejados como pudiera parecer. Con ligeras discrepancias, los diferentes métodos nos llevan a soluciones bastante similares, y de confirmarse este tipo de resultados en el análisis de otros casos, la elección de uno u otro podría basarse más en criterios de operatividad o preferencias personales que en la pretendida superioridad de algunos de ellos.

Por ejemplo, al elegir los métodos basados en la TRI, podríamos tener presentes los inconvenientes prácticos que se asocian a éstos, pues precisan muestras relativamente amplias para conseguir una suficiente bondad de ajuste en la estimación de parámetros, y requieren de cálculos complejos y procedimientos sofisticados. Ambas razones podrían limitar su utilización frente a los restantes métodos, más intuitivos y con menos exigencias en cuanto a tamaño muestral o esfuerzo de cálculo.

A pesar de todo, aunque los métodos coincidan en señalar cuáles son los ítems que presentan DIF, es difícil interpretar las razones del mismo. Todos estos métodos cuentan con el inconveniente de no ofrecer pistas para la interpretación de las causas que originan el posible sesgo. De ahí que un posterior análisis lógico de la redacción y el contenido del ítem sean imprescindibles de cara a mejorar el reactivo o decidir su exclusión del instrumento.

REFERENCIAS BIBLIOGRÁFICAS

- Angoff, W.H. (1982). Uses of difficulty and discrimination indices for detecting item bias, en R.A. Berk (Ed.) *Handbook of methods for detecting item bias*. Baltimore, John Hopkins University Press.
- Angoff, W.H. y Ford, S.F. (1973). Item-race interaction on a test of scholastic aptitude, en *Journal of Educational Measurement*, 10, 95-105.
- Assesment Systems Corporations (1995). *XCALIBRE. Marginal maximum likelihood estimation program for the 2-and 3-parameter IRT model. Version 1.00*. St. Paul, MN: Autor.
- Bock, R.D. y Lieberman, M. (1970). Fitting a response model to n dichotomously scored items, en *Psychometrika*, 35, 179-197.
- Camilli, G. (1979). *A critique of the chi-square method of assessing item bias*. University of Colorado, Laboratory of Educational Research.
- Camilli, G. y Sephard, L.A. (1994). *Methods for identifying biased test items*. Thousand Oaks, California, Sage Publications.
- Cohen, A.S. y Kim, S.H. (1993). A comparison of Lord's chi-square and Raju's area measures in detection of DIF, en *Applied Psychological Measurement*, 17, 39-52.
- Diamond, J.J. (1992). A Graphic Procedure for Studying Differential Item Functioning, en *Journal of Experimental Education*, 60 (4), 351-57.
- Dorans, N.J. y Holland, P.W. (1993). DIF detection and description: Mantel-Haenszel and standardization. En P.W. Holland y H. Wainer (Eds.), *Differential item functioning*. Hillsdale, New Jersey, Lawrence Erlbaum.
- Dorans, M.J. y Kurlick, E.M. (1983). *Assessing unexpected differential item performance of female candidates on SAT and TSWE forms administered in December 1977: an application of standardization approach*. Princeton, New Jersey, Educational Testing Service.
- Dorans, M.J. y Kurlick, E.M. (1986). Demonstrating the utility of the standardization approach to assessing unespected differential item performance on the Scholastic Aptitude Test, en *Journal of Educational Measurement*, 23, 355-368.
- García, E. (Dir.) (1993). *Los Tests Adaptativos Computerizados (TACs): su aplicación al Area de Geografía, Historia y Ciencias Sociales en la Educación Secundaria Obligatoria*. Proyecto aprobado por el CIDE en el Concurso Nacional de Proyectos de Investigación. Convocatoria 1993.
- Hambleton, R.K. y otros (1993). Advances in the detection of differentially functioning test items, en *European Journal of Psychological Assessment*, 9 (1), 1-18.
- Hanson, B.A. y Feinstein, Z.S. (1995). A Polynomial Loglinear Model for Assessing Differential Item Functioning. Comunicación presentada al *Annual Meeting of the American Educational Research Association*. San Francisco, CA.

- Kim, S.H. y Cohen, A.S. (1992). Effects of linking methods on detection of DIF, en *Journal of Educational Measurement*, 29(1) 51-66.
- Lord (1980). *Applications of item response theory to practical testing problems*. Hillsdale, Nueva Jersey, LEA.
- Mazor, K.M. y otros (1994). Identification of Nonuniform Differential Item Functioning Using a Variation of the Mantel-Haenszel Procedure, en *Educational and Psychological Measurement*, 54 (2), 284-91.
- Mazor, K.M. y otros (1995). Using Logistic Regression and the Mantel-Haenszel with Multiple Ability Estimates to Detect Differential Item Functioning, en *Journal of Educational Measurement*, 32 (2), 131-144.
- Miller, T.R. y Spray, J.A. (1993). Logistic Discriminant Function Analysis for DIF Identification of Polytomously Scored Items, en *Journal of Educational Measurement*, 30 (2), 107-22.
- Muñiz, J. (1992). *Teoría clásica de los tests*. Madrid, Pirámide.
- Muñiz, J. y Hambleton, R.K. (1992). Medio siglo de Teoría de Respuesta a los Ítems, en *Anuario de Psicología*, 52, 41-66.
- Nandakumar, R. (1994). Development of a Valid Subtest for Assessment of DIF/Bias. Comunicación presentada al *Annual Meeting of the American Educational Research Association*. New Orleans, LA.
- Oshima, T.C. y otros (1994). Differential Item Functioning for a Test with a Cutoff Score: Use of Limited Closed-Interval Measures, en *Applied Measurement in Education*, 7 (3), 195-209.
- Pang, X. y otros (1994). Performance of Mantel-Haenszel and Logistic Regression DIF Procedures over Replications Using Real Data. Comunicación presentada al *Annual Meeting of the American Educational Research Association*. New Orleans, LA.
- Zwick y otros (1994a). Assessment of Differential Item Functioning for Performance Tasks, en *Journal of Educational Measurement*; 30 (3), 233-51.
- Zwick, R. y otros (1994b). A Simulation Study of Methods for Assessing Differential Item Functioning in Computerized Adaptive Tests, en *Applied Psychological Measurement*, 18 (2), 121-140.

Autora: M^a del Consuelo Vélaz de Medrano Ureta
Dirección: C/ Barquillo 15 B, 2º E; 28004 Madrid; Tfno.: 91 5225602
Director: Dr. D. Arturo de la Orden Hoz
Dpto.: Métodos de Investigación y Diagnóstico en Educación
Dirección: Facultad de Educación. Universidad Complutense de Madrid.
 Edificio «La Almudena». Ciudad Universitaria.

Descriptores

Investigación educativa; Creencias sobre la ciencia; Hábitos científicos; Política científica en educación; Didáctica de la ciencia; Epistemología.

Bibliografía fundamental

C.I.D.E. (1995). *Doce años de investigación educativa. Catálogo 1983-1994*. Madrid: Servicio de Publicaciones del Ministerio de Educación y Ciencia.
 Cohen, I.B. (ed.) (1993). *The natural sciences and the social science: some critical and historical perspectives*. Dordrecht: Kluwer Academic Publishers.
 Feyerabend, P. (1986). *Tratado contra el método. Esquema de una teoría anarquista del conocimiento*. Madrid: Tecnos.
 Koulaidis, V. (1987). *Philosophy of Science in relation to curricular and pedagogical issues. A study of science teachers' opinions and their implications*. London: Institute of Education, University of London. Tesis inédita.
 Kuhn, T.S. (1975). *La estructura de las revoluciones científicas*. Madrid: Fondo de Cultura Económica.
 Lakatos, I. (1989). *La metodología de los programas de investigación científica*. Madrid: Alianza.
 Ministerio de Educación y Ciencia (1989). *Plan de Investigación Educativa y de Formación del Profesorado*. Madrid: Servicio de Publicaciones del M.E.C.
 Novak, J.D. y Gowin, D.B. (1988). *Aprendiendo a aprender*. Barcelona: Martínez Roca.
 O.C.D.E. (1995). *Educational research and development. Trends, issues and challenges*. París: O.C.D.E.
 Popper, K. (1985). *La lógica de la investigación científica*. Madrid: Tecnos.
 Ryan, A., Aikenhead, G. (1992). The development of a new instrument: «Views on Science-Technology-Society» (VOSTS), *Science Education*, 76 (5), pp. 477-491.

Problema de investigación (planteamiento e hipótesis de trabajo)

Definición del problema central: «Tomando como marco de referencia los planteamientos de *cuatro sistemas epistemológicos contemporáneos (Inductivismo, Racionalismo Crítico, Contextualismo y Relativismo) ¿qué concepción de la ciencia predomina entre los investigadores en ciencias de la educación?*». Al mismo tiempo se ha tratado de dar respuesta a otros elementos consustanciales al problema (siguiendo el criterio de Lawson y otros (1989), el lugar de las hipótesis lo ocupan aquí los objetivos de investigación): a) Caracterizar la muestra de investigadores: Edad; género; formación inicial; actividad profesional; experiencia investigadora; grado de definición, consistencia y especialización de su práctica científica; modalidad practicada (básica/aplicada/ambas); metodología; forma de investigar (en solitario/en equipo estable/en equipos coyunturales); formación en Epistemología; b) Analizar si las creencias sobre la ciencia varían en función de las variables descriptivas; c) Conocer la opinión de los investigadores con respecto a una serie de aspectos de la política científica en educación (rasgos a valorar positivamente en un proyecto de investigación para su financiación pública; factores más influyentes en el progreso de un campo de investigación en educación; ventajas e inconvenientes de las convocatorias de ayudas a la investigación promovidas por el M.E.C. a través del C.I.D.E. para

Autora: M^a del Consuelo Vélaz de Medrano Ureta
IMAGEN DE LA CIENCIA, PRÁCTICAS Y HÁBITOS CIENTÍFICOS DE LOS INVESTIGADORES EN CIENCIAS DE LA EDUCACIÓN

CAT ^(a)	ANO ^(a)	CLASIFICACIÓN ^(a)
D	1 9 9 6	5 8 9 9 -
N.º CITAS ^(b) N.º PÁGINAS		
7	4 1 0 9 8 6	6

incentivar la investigación educativa de calidad); d) Ofrecer unas estrategias didácticas orientadas a promover y facilitar un proceso de enseñanza-aprendizaje más crítico del significado y características de la actividad científica.

Muestra y método de muestreo

La población viene definida por el conjunto de investigadores registrados en la base de datos del Centro de Investigación y Documentación Educativa (C.I.D.E.) en el período 1985-1995 (N=409). El ámbito de aplicación del estudio ha sido, pues, nacional. El muestreo realizado fue aleatorio, estratificado (con afijación proporcional en función de una variable geográfica con dos niveles: Madrid y resto del Estado) y en dos etapas. Como resultado se obtuvo una muestra de 300 investigadores. El índice de respuesta al cuestionario fue del 38,3% (106 investigadores), lo que puede considerarse satisfactorio dado que en este tipo de estudios por correo no suele superar el 20%.

Metodología del trabajo

En el diseño de investigación se distinguen tres fases: I) *Estudio comparado* de cuatro concepciones alternativas de la ciencia (Inductivista, Racionalista-Crítica, Contextualista y Relativista) e identificación de los temas clave en torno a los que éstas se pronuncian; II) *Proceso de definición y operativización del constructo a explorar*: a) Identificación de la estructura argumental subyacente a las distintas posiciones sobre la ciencia: las dimensiones del dominio; b) Elaboración de un sistema de indicadores —operativizados en variables— que proporcione una representación global válida y permita elaborar un instrumento para analizar tendencias; c) Organización y jerarquización de los conceptos clave a explorar y de sus interrelaciones (Redes Sistémicas, Mapas Proposicionales y Representaciones generales con una utilidad didáctica adicional); III) *Estudio empírico descriptivo-exploratorio*: elaboración y aplicación de un cuestionario sobre la imagen de la ciencia y las prácticas científicas de los investigadores.

Técnicas de análisis

- Técnicas empleadas para la validación del instrumento: a) Fiabilidad como consistencia interna mediante el *Coefficiente Alpha de Cronbach*. Se emplea el *Coefficiente de Correlación Cuádruple Φ* para la estimación de la asociación de variables dicotómicas en un conjunto reducido de ítems; b) Validez formal y de contenido: mediante *sistema de jueces*.
- Técnicas para el análisis descriptivo de la muestra: Índices porcentuales, medidas de tendencia central y de dispersión; Una técnica no paramétrica para estimar la significatividad de las relaciones entre variables descriptivas: el *Estadístico Ji Cuadrado* como prueba de independencia (estimando la magnitud de las asociaciones significativas con el *Coefficiente Φ de Cramer*).
- Técnicas para la exploración de la imagen de ciencia predominante: a) *Análisis Cluster de Variables* (Jerárquico, Aglomerativo y con el Método de promedio entre grupos); b) *Análisis Cluster de Sujetos* (combinado con el uso del *Estadístico Ji Cuadrado* como prueba de independencia).

Conclusiones

Los investigadores de la muestra parecen tener una opinión bastante definida acerca del conocimiento y de la actividad científica, pudiendo distinguirse tres imágenes diferenciadas de la ciencia. La imagen predominante (entre más del 45% de los investigadores) está muy próxima a las corrientes epistemológicas actuales que introducen una dimensión psicológica, histórica y sociológica en la lógica de la investigación científica. Asimismo las prácticas científicas de quienes la sostienen parecen ajustarse a los rasgos que dichas corrientes atribuyen a los períodos de madurez de un campo científico (constancia en el objeto de estudio, referente teórico bien definido, y trabajo en un equipo estable de investigación). Pese a ello, el predominio de muchos de los supuestos y prácticas positivistas radicales en los otros dos grupos (especialmente en el constituido por el 31% de la muestra) hacen pensar que este enfoque sigue teniendo una presencia muy importante. Los resultados obtenidos permiten concluir que se ha realizado un importante avance —que debe continuar— con respecto a estudios precedentes en: a) La definición y operativización del dominio a explorar, destacando el «sistema de indicadores» y un cuestionario con altos índices de consistencia interna; b) La mejora del conocimiento del perfil académico, profesional y científico de los investigadores en temas educativos en nuestro país, así como de sus opiniones y valoraciones sobre algunos aspectos de política científica en investigación educativa.

y el volumen de investigación desarrollado sobre el mismo, se ha aconsejado su estudio y adaptación a nuestro ámbito cultural.

Al proceso de adaptación realizado se han añadido dos estudios de gran interés y que no existen en la edición original del KM-R:

a) una fundamentación teórica de la prueba basada en algunas de las conclusiones aportadas por el enfoque del Procesamiento de la Información sobre los procesos de aprendizaje de la matemática,

b) una revisión lógica de la Validez de Contenido basada en dos estudios de adecuación curricular: un análisis de la correspondencia de los dominios del KM-R con el currículum del NCTM de 1989, según la estrategia de Romberg (1992), y la determinación de la existencia de cada dominio de la Escala en el Currículum LOGSE y el momento de su consecución según el criterio de un grupo de expertos.

Muestra y método de muestreo

La **muestra** aleatoria de sujetos (N=1.238) es representativa de la población escolar de la Comunidad Valenciana (Preescolar, EGB-LOGSE y 1º BUP-FP), estratificada por curso o nivel de estudios, sexo, tipo de colegio, clase social, lengua, etc.

Metodología del trabajo

Adaptación socio-pedagógica y curricular; análisis psicométrico y diferencial.

Técnicas de análisis

Análisis de ítems. Revisiones lógica y empírica. Análisis de la Fiabilidad (mitades y consistencia interna). Evidencias de Validación. Análisis Factorial; Análisis Cluster y estudios diferenciales de grupos de rendimiento; validación de contenido mediante estudios basados en juicio.

Conclusiones

Como conclusión del trabajo de adaptación realizado, se puede afirmar que la Escala *KeyMath-Revised* de A. Connolly puede ser considerada una prueba general de matemáticas elementales referida al currículum y con una organización interna correspondiente al currículum. Ha mostrado unos niveles de bondad métrica muy aceptables, mejorando los estudios originales de Connolly, que avalan su uso; asimismo, teniendo en cuenta los resultados obtenidos en los diferentes estudios de validación, se revela como un instrumento diagnóstico de gran utilidad.

Además, la fundamentación teórica realizada a partir de algunas conclusiones del enfoque del Procesamiento de la Información y las revisiones lógicas de la Validez de Contenido realizadas, no existentes en la adaptación original de Connolly, otorgan al KM-R una mayor riqueza de interpretación diagnóstica y permiten diseñar programas de intervención basados en el propio currículum y en las conclusiones de la propia psicología cognitiva.

En el estudio han participado 18 profesores/as de Educación Primaria, de los cuales, 6 impartían su docencia en el primer curso, y 12 lo hacían en el segundo.

Todos llevaban más de 12 años de servicio, y 10 de ellos, estaban estudiando o habían finalizado los estudios de Pedagogía, por lo que el interés por el trabajo propuesto era bastante alto.

Todos los colegios pertenecían a la comarca escolar de Granada.

El interés de la muestra no es representara una población determinada, sino obtener la mayor cantidad de información posible, para obtener teoría de la propia práctica.

Metodología de trabajo

Para dar respuesta a los objetivos planteados, hemos recogido información por dos vías diferentes:

a) A través de una escala de observación elaborada por nosotros.

b) A través de los diarios de los observadores.

El proceso seguido para la elaboración de la escala de observación ha sido el siguiente:

- 1.— Análisis del Diseño Curricular par Educación Primaria de la Junta de Andalucía.
- 2.— Selección de categorías.
- 3.— Observación directa de la actuación de los profesores en el aula, para seleccionar los rasgos que van a definir cada una de las categorías.
- 4.— Adscripción de los rasgos a cada una de las categorías.

La finalidad de utilizar los diarios de los observadores fue recoger la misma información por dos vías diferentes.

Técnicas de análisis

Hemos realizado análisis cualitativo, entendiendo por tal, el proceso de extraer sentido a los datos recogidos. Hemos aplicado el programa AQUAD, 3.0 (Huber, 1988). Las fases que hemos seguido las podemos resumir:

- 1.— Codificación de los textos.
- 2.— Descripción de los códigos que utiliza cada profesor en las diferentes materias en que ha sido observado.
- 3.— Representación gráfica de los datos, indicando la frecuencia de aparición de los mismos.
- 4.— Comprobación de hipótesis cualitativas.

Los datos han sido tratados con el paquete estadístico AQUAD 3.0.

Conclusiones

- Del análisis de los datos obtenidos hemos podido constatar cuales son las actuaciones de los profesores en las diferentes materias (Matemáticas, Conocimiento del Medio, Lenguaje, Expresión Plástica) las cuales no vamos a reflejar por falta de espacio.
- Los profesores que han participado en el trabajo, cuando imparten su docencia, actúan de acuerdo con las ideas constructivistas de las que partíamos al enmarcar teóricamente nuestro trabajo.
- No existen diferencias en cuanto a la actuación de los profesores en las materias de Lenguaje, Matemáticas y Conocimiento de Medio, aunque si difieren en Expresión Plástica.
- Una de las implicaciones mas importantes del trabajo consideramos que es la elaboración y entrega, de un informe personal a cada uno de los participantes, el cual les ha ayudado a conocer mejor su propia actuación en el aula para así mejorar aquéllo que estimen oportuno.

La población la constituyen todos los centros de primaria y secundaria de Cataluña.

La muestra es estratificada de acuerdo con las siguientes variables:

- Titularidad
- Nivel educativo
- Número de alumnos
- Tipo de población
- Provincia

A cada uno de los 50 centros se pasó el cuestionario a 6 directivos, 6 profesores, 2 padres/madres y 2 alumnos.

Además, y como contraste, se pasó a los asistentes a las IV Jornadas de Dirección Escolar que organizó el FEAEC.

Metodología del trabajo

Tratándose de una investigación descriptiva se construye un cuestionario, de acuerdo con las siguientes fases:

- Delimitación operativa de las variables de estudio.
- Elaboración de un cuestionario abierto y debate por el equipo investigador.
- Consulta a diferentes expertos y jueces.
- Elaboración del cuestionario provisional y validación por el equipo de investigación.
- Validación por consulta a representantes de las categorías destinatarias (maestros y madres/padres).

La aplicación se realizó por medio de encuestadores que estuvieron presentes mientras se respondía al cuestionario.

Técnicas de análisis

Para el tratamiento informático se ha utilizado el paquete SPSSX en la versión "release" 3.1. Concretamente los procedimientos:

MEANS para el pre-análisis de las diferencias entre medias de las diversas submuestras.

T-TEST para analizar la igualdad o desigualdad de las medias y desviaciones típicas entre subgrupos independientes.

Conclusiones

Se han establecido dos tipos de conclusiones: respeto a las temáticas tratadas y respeto a los colectivos consultados.

Por lo que se refiere a las modificaciones que la LOPEGCD introduce en relación a la dirección de los centros educativos hay que señalar que:

- El tiempo de ejercicio del cargo debe ser superior a 3 años. Cuando no hay candidatos del centro, se considera que el período de ejercicio debe ser inferior.
- En relación a la obligatoriedad de colaborar en la evaluación, las respuestas varían. Es mucho más aceptada la evaluación de centros, seguida de los directivos y, por último, la del profesorado.
- La necesidad de formación específica para el cargo (alrededor de 200 horas) y la elaboración de un proyecto de dirección es comúnmente aceptada.

Por lo que se refiere al modelo de dirección:

- La elección del director/a está bien vista por todos los colectivos al igual que están de acuerdo en que el Consejo Escolar sea el órgano máximo de gobierno.

Respeto a los colectivos cabe señalar que las respuestas de las personas más preocupadas y formadas en temas de organización y dirección de centros manifiesta de forma significativa la necesidad de mayor formación para el cargo, mayores atribuciones, proceso de selección restringido y período de ejercicio del cargo más largo.

Autor: Jesús Domingo Segovia
Dirección: C/ Santa Rosa, 6-1º B
 18200 Maracena (Granada)
Director: Antonio Bolívar Botía
Dpto.: Didáctica y Organización Escolar
Dirección: Campus Universitario de Cartuja
 18071 Granada
Centro: Facultad de Ciencias de la Educación. Universidad de Granada

Descriptores

Ciclo de vida, ciclo de desarrollo, historia de vida, método biográfico, desarrollo organizativo, cultura, memoria institucional, procesos de desarrollo y mejora, cambio, aprendizaje organizacional, mejora de la escuela...

Bibliografía fundamental

Adizes, I. (1993). *Ciclos de vida de la organización: cómo y por qué crecen y mueren las organizaciones y qué hacer al respecto*. Madrid, Ed. Díaz Santos.

Bertaux, D. (Ed.) (1983). *Biography and society: The life history approach in the social sciences*. London, Sage.

Connelly, F.M. & Clandinin, D.J. (1990). «Stories of experience and narrative inquiry», *Educational Researcher*, 19(5), 2-14.

Fernández, M. (1994). *Una aproximación biográfica al desarrollo profesional de maestros y maestras de educación infantil: ciclo vital, identidad, conocimiento y cultura*. Universidad de Granada, Tesis doctoral (inédita).

Gómez-Pallete, F. (1995). *La evolución de las organizaciones*. Madrid, Noesis.

Hüber, G.L. (1991). *Análisis de datos cualitativos con ordenadores: principios y manual del paquete de programas AQUAD 3.0*. Sevilla, Ed. Carlos Marcelo.

Kelchtermans, G. y Vandenbergh, R. (1994). «Teachers' professional development: a biographical perspective», *Journal of Curriculum Studies*, 26(1), 45-62.

Marinas, J.M. y Santamarina, C. (Eds.) (1993): *La historia oral: métodos y experiencias*. Madrid, Debate.

Power, C.F., Higgins, A. & Kohlberg, L. (1989). *Lawrence Kohlberg, approach to moral education*. New York, Columbia University Press.

Pujadas, J. (1992). *El método biográfico: el uso de las historias de vida en ciencias sociales*. Madrid, CIS.

Polhinghome, D.E. (1995). *Qualitative Studies in Education*, 8(1), 5-24.

Ricoeur, P. (1987). *Tiempo y narración (2 vols.)*. Madrid, Cristiandad.

Van de Ven, A.H. & Poole, M.S. (1995). «Explaining development and change in organizations», *Academy of Management Review*, 20(3), 510-540.

Problema de investigación

Las escuelas son realidades socialmente construidas; por ello, cada una tiene su propia cultura con una memoria institucional subjetivamente vivenciada y que se desarrolla y se va dotando de identidad a lo largo de un ciclo de vida, en el que se encuentra sentido y continuidad cada cambio, momento, acción, impacto, resistencia, etc. Este proceso de desarrollo cultural e histórico es mucho más que una simple evolución o un cúmulo de dinámicas de cambio sin sentido de continuidad.

Es relevante el estudiar la historia de un centro (estudio de caso) desde la reconstrucción biográfico-narrativa de su ciclo de vida, por ofrecernos la posibilidad de recoger el saber adquirido, dinamizar y cuestionar lo que se hace, dentro del desarrollo organizativo del centro, en términos singificativos actuales para los propios miembros del colegio —surgidos desde la memoria institucional y como adelanto de hacia dónde caminamos— y describir comprensivamente qué elementos se han institucionalizado y por qué, como claves para interpretar la situación actual y proyectar el futuro. Además de estas claves descriptivas y comprensivas del proceso de reconstrucción cultural y de desarrollo del centro, caracterizamos el ciclo de vida del centro educativo objeto de estudio.

Muestra y método de muestreo

Se trata de un estudio de caso, centrado en un centro concreto. La muestra ha sido el propio centro: sus componentes personales (padres y profesores) y una selección de informantes claves —surgidos desde los contactos iniciales y las contextualizaciones y estudios colaterales iniciales—, a los que se han ido añadiendo cuantos casos hemos estimado oportunos en función del proceso de bola de nieve que iniciamos para lograr la saturación de las informaciones y de los elementos coincidentes en base a las comparaciones y yuxtaposiciones de historias y discusiones en grupo, al timpo que se hacía una búsqueda sistemática de casos negativos o puntos oscuros cuya clarificación pudiera contribuir a completar y mejorar la historia. Hemos contado con 35 importantes más los participantes a las discusiones en grupo de los tres sectores.

Autor: Jesús Domingo Segovia
**CICLO DE DESARROLLO DE UN CENTRO EDUCATIVO:
 APROXIMACIÓN BIOGRÁFICO-INSTITUCIONAL**

CAT ^(a)		AÑO ^(a)		CLASIFICACIÓN ^(a)	
D	1	9	9	6	
					5
					8
					9
					9
N.º CITAS ^(b)		N.º PÁGINAS			
7	3	6	1	7	8
					5

Metodología de trabajo

Combina los dos métodos de estudio cualitativo sobre narrativa (Polkinghorne, 1995): investigación narrativa y de narrativas. La aproximación biográfico-institucional se concreta en la construcción dialéctica de un nuevo relato polifónico del proceso (Lewis, 1964; Pujadas, 1992). Esta historia de vida se va articulando a partir de una genealogía del contexto o biograma —como reconstrucción histórica de su proceso de desarrollo— a partir múltiples relatos narrativos individuales de vida del centro y las propias contextualizaciones y tomas de contacto realizadas mediante el estudio de diversas fuentes documentales escritas del centro —*fase de aproximación y toma de contacto, contextualización inicial*—. Seguidamente, una vez obtenidos los temas centrales o generadores y la estructura temporal y conceptual sobre la que ir armando la historia, iniciamos un proceso de profundización y comprensión sucesiva —adaptando el modelo de Kelchtermans y Vandenberghe (1994): a) *Aproximación* (autoinformes personales paralelos), para recopilar elementos tangibles e intangibles de la organización, significados, relaciones, momentos críticos, lograr una trama y detectar puntos oscuros y nuevos informantes; b) *Profundización* (diálogos en profundidad), construcción de una identidad narrativa y una reflexión amplia y concreta por aspectos, momentos, etc.; y c) *Validación y reconstrucción de un relato polifónico de vida institucional* (entrevistas y discusiones de grupo) confrontar, consensuar, validar y completar)

Técnicas de análisis

El proceso ha sido complejo, con variadas técnicas en función de cada momento del trabajo, objetivos, contenidos e instrumentos que habíamos seleccionado en cada una de ellas. Se ha utilizado un modelo cíclico de comprensión de datos similar al empleado por Miles y Huberman (1984) con diversos momentos de análisis y de comparación interfuentes e historias (Parrilla, 1990) e inspirado en la inducción analítica (Goetz y LeCompte, 1988) y en las fases del método biográfico (Bertaux, 1981). Para llegar a una triangulación interactiva de las diversas historias y/o evidencias referidas a un mismo tema, hasta llegar a un nuevo relato polifónico comprensivo (consensuado dialécticamente) del proceso de desarrollo seguido.

El proceso de análisis sigue las fases de: codificación y categorización (reducción de datos o descontextualización), recontextualización y categorización —fichas de metáforas y secuencias, códigos (tres modelos) y dimensiones de análisis— y determinación de vínculos y relaciones. El proceso en sí consta de: (1) un análisis de contenido de los documentos objetos de estudio; (2) seguidamente, en la fase de aproximación, se hizo un análisis de relatos (mapas y esquemas de ideas, matrices relacionales y aplicación del programa AQUAD 3.0 —vínculos y relaciones entre categorías, secuencias, frecuencias de aparición...— y validación de informes de progreso; (3) seguidamente —con las entrevistas— se repitió el proceso y se aplicó el programa AQUAD 3.0 (componente de minimalización: caracterización de grupos, fases y momentos, implicantes esenciales y condiciones negativas) y, por último, (4) se sometió dialécticamente a validación y aprobación personal (cada informante) y grupal de las conclusiones y datos obtenidos. En este sentido se cuidó sobremanera, aparte de los criterios propios de la investigación cualitativa, la implicación personal del narrador e investigador, la presunción de veracidad (identidad narrativa), las contextualizaciones constantes y las validaciones dialécticas sucesivas y procesuales que anulan los efectos del investigador nativo y las dificultades propias de propio enfoque.

Conclusiones

Un centro escolar, como cultura en desarrollo, es mucho más que su estructura, artefactos o materializaciones y productos; es dinámico, se reconstruye constatemente a lo largo de un ciclo de vida —aunque muy variable, personalizado y con características especiales/específicas—. Éste puede ajustarse al modelo de ciclo de vida y cambio propuesto para las organizaciones en general (Adizes, 1993; Gómez-Pallete, 1995) y apunta recurrentemente a seguir las fases indicadas para todo cambio educativo (Escudero, 1991, 1993). En él, se van asentando las propias dinámicas de reflexión, análisis e integración del propio cambio, desde una perspectiva de centro y con distintos ritmos, contenidos, conflictos, enfoques, estilos de liderazgo, etc. en función de cada fase de desarrollo de la organización y de las personas, componentes y circunstancias que rodean la escuela en cada caso. Al tiempo que se ha realizado una descripción comprensiva del proceso del desarrollo seguido por el centro objeto de estudio.

La comprensión y análisis de las organizaciones escolares, en cuanto culturas que forman un sistema complejo con profundas interrelaciones, demanda aproximaciones complejas que incluyan a los propios agentes, sus esquemas conceptuales, memoria colectiva, sistema de significados, etc. (como sería una aproximación biográfica), que consideren la integración e interacción de múltiples percepciones que le aportan identidad y comprensión escénica (individual-colectiva, subjetiva-polifónica, formal-ideográfica, diacrónica-sincrónica, prospectiva-retrospectiva, manifiesta-oculta, acciones-silencios, valores y contextos, etc.).

El modelo de ciclo de vida, aún con limitaciones internas para los centros escolares, se constituye como un marco conceptual interesante —próximo a los planteamientos de los modelos de proceso y de mejora de la escuela— para estudiar el desarrollo de una cultura u organización. La identificación de los momentos críticos, el cúmulo complejo de circunstancias, roles, dificultades y apoyos, azar, constancia, etc. son indicadores potentes para detectar en qué momento de desarrollo se encuentra la organización, cuál es su tipología cultural y qué hacer en cada caso o momento para potenciar su crecimiento; puesto que lo que es apropiado a un determinado período, frecuente puede no ser adecuado para otro.

Los procesos de cambio, para alcanzar madurez, necesitan de un tiempo de gestación, crecimiento e institucionalización dilatado en el tiempo y han de implicar al centro en su conjunto para ser significativos.

IX CONGRESO NACIONAL DE MODELOS DE INVESTIGACIÓN EDUCATIVA

***Nuevas realidades educativas,
nuevas necesidades metodológicas***

Málaga, 21, 22 y 23 de Octubre de 1999

INFORMACIÓN:

Universidad de Málaga
Facultad de Ciencias de la Educación
Área de Métodos de Investigación y Diagnóstico en Educación
Campus de Teatinos
29071 Málaga
Teléf. 95 213 10 74
Fax 95 213 26 26
E-mail: angulo@uma.es

ÁREA MIDE

Información

Durante los días 24 al 26 de junio de 1998 se ha celebrado en Tenerife la III Reunión Estatal del Área de Métodos de Investigación y Diagnóstico en Educación. A dicha reunión asistieron representantes de las Universidades de Alicante, Barcelona, Cádiz, Cantabria, La Coruña, Granada, La Laguna, Las Palmas, Madrid, Málaga, Murcia, Oviedo, País Vasco, Salamanca, Santiago, Sevilla, UNED, Valencia y Valladolid. En este encuentro se analizó la información recogida entre el profesorado de las distintas universidades españolas sobre metodologías de enseñanza y formas de evaluación empleadas en las distintas asignaturas. Este Seminario es continuación de los dos anteriores celebrados respectivamente, en Salamanca (junio 1993), dónde se analizaron los contenidos de las distintas materias y asignaturas que el Área imparte en las distintas titulaciones. En el segundo Seminario, celebrado en La Coruña (junio 1996), se abordó el practicum de las diferentes titulaciones en las que el Área también está presente.

Lo que surgió como interés de unos pocos profesionales, se ha convertido en necesidad de muchos, pues cada vez ha sido mayor el profesorado y el número de universidades que han participado en dichos encuentros. Estas reuniones reflejan la necesidad no sólo de compartir preocupaciones sino de aunar esfuerzos por mejorar la calidad de la docencia en las asignaturas que imparte el profesorado del Área MIDE y en su interés por formar buenos profesionales.

Un objetivo de estas jornadas es la coordinación y el intercambio profesional entre las diferentes universidades para poder analizar las similitudes y peculiaridades de las distintas titulaciones en cada una de las comunidades autónomas.

Es un momento idóneo para reflexionar y establecer orientaciones mínimas que puedan guiar las líneas de actuación para la necesaria reformulación de los planes de estudio.

El trabajo que se ha ido realizando en estas reuniones tendrá su continuación en la Universidad de Murcia, Departamento M.I.D.E., en septiembre del año 2000.