

## **LA PRODUCTIVIDAD DE LA INVESTIGACIÓN SOBRE EVALUACIÓN DE PROGRAMAS EDUCATIVOS ESPAÑÓLES (1975-2000)**

Jorge Expósito López  
Antonio Fernández Cano\*  
Dpto. MIDE  
Universidad de Granada

### **RESUMEN**

*Este estudio es una revisión cuantitativa que trata de sintetizar, en cuanto indicadores de productividad, las realizaciones de la investigación educativa española sobre evaluación de programas durante el periodo 1975-2000.*

*Partiendo fundamentalmente de una búsqueda en la base de datos del I.S.O.C. (Instituto de Información y Documentación en Ciencias Sociales, dependiente del C.S.I.C.) complementada con búsquedas cruzadas manuales se realiza la revisión de 138 estudios publicados como artículos de investigación en revistas científicas. Una serie de indicadores, compilados podrían ayudar a inferir ciertas consideraciones sobre evaluación de programas en educación en este campo y periodo.*

### **ABSTRACT**

*This study is a scientometric review trying to synthesise, according to productivity indicators, the primary Spanish studies available about program evaluation into the educational field during the period 1975-2000.*

*Beginning from a literature retrieval in the ISOC / CSIC base data and completed with other cross-references retrieved, a review is accomplished with 138 studies published as research papers in educational scientific journals. A series of compiled indicators could help to infer some considerations about the evaluation of educational programs in this field and time.*

---

\* Dpto. Métodos de Investigación y Diagnóstico en Educación. Facultad de Ciencias de la Educación. Universidad de Granada (UGR). Campus de Cartuja, s/n. 18071 Granada. Teléfonos 3458 24 42 53 y 3458 74 10 06. Fax 3458 24 99 71. E-mail: [afcano@platon.ugr.es](mailto:afcano@platon.ugr.es)

## DESCRIPTORES

Evaluación de Programas Educativos, Investigación Española, Estudio Científico, Productividad.

### I. INTRODUCCIÓN

En 1929, Ortega y Gasset escribía en su libro *La rebelión de las masas* que: «Esta [la ciencia] necesita de tiempo en tiempo, como orgánica regulación de su incremento, una labor de reconstitución, y..., esto requiere un esfuerzo de unificación, cada vez más difícil que complica regiones más vastas del saber total».

Esa labor de reconstitución pasa básicamente por dos caminos: La investigación de síntesis, a la búsqueda de la verdad definitiva por inducción, y la propia evaluación de la investigación, como medio de denotar la validez y fecundidad de los hallazgos, primarios o secundarios, que se infieren (Fernández Cano, 1995). A medio camino de ambas vías surge la metodología científimétrica con profundas implicaciones para la síntesis formal, que no de contenido, y con pautas de racionalidad paraevaluativa denotadas por el ajuste a patrones propios de la ciencia.

El campo de estudio a indagar aquí es el de la evaluación de programas educativos. Por tanto, cuando nos interrogamos sobre evaluación estamos en cierta medida realizando metaevaluación (lo que hay más allá de la evaluación, en analogía al planteamiento terminológico aristotélico de la metafísica).

La pertinencia de estudios metaevaluativos en educación ha sido altamente señalada e indicada por autores extranjeros (Scriven, 1974; Stufflebean, 1978, 1981; Shadish, Cook y Leviton, 1991) y nacionales (Bartolomé, 1990; De la Orden y Martínez de Toda, 1992 o Fernández Cano y Bueno, 1998). Sin embargo, son pocos los estudios de evaluación que son evaluados, pues existe una resistencia a ser evaluado que pone en tela de juicio esta práctica (Worthen y Sanders, 1987). Tal meta-evaluación admite distintas aproximaciones, una de ellas es la que aquí usamos utilizando una metodología científimétrica.

La National Development and Career Association (N.D.C.A.), asociación de profesionales de la orientación, posee un servicio que anualmente revisa los trabajos de intervención e investigación en esta área. En nuestro país, aunque no admite comparación, se comienzan a estructurar algunas bases de datos como Isis o compendios bibliográficos como Bordón, y a realizar revisiones sobre programas específicos (Salvador y Peiró, 1986; Rodríguez, 1993; Repetto, 1995).

Sin embargo, las revisiones aparecidas en la década de los 90 sobre evaluación de programas siguen siendo relativamente limitadas (Arruabarrena y Paul, 1994; Repetto, 1993 y 1995) y referidas en su mayor parte a lo producido en el ámbito anglosajón. Como afirman Álvarez Rojo y Hernández (1997) «no estamos sobrados en nuestro país de trabajos que lleven a cabo revisiones sistemáticas», y también «una revisión consiste en esencia en el análisis de los trabajos publicados (libros, artículos y otros materiales) sobre un determinado tema en un periodo de tiempo dado con objeto de determinar la

estructura subyacente y las líneas de desarrollo de ese tópico en ese lapsus temporal» (p. 80).

Por otra parte, se ha comenzado a producir importante literatura en España (Gelpi, 1987; Repetto, 1987; Colás y Rebollo, 1993; Medina y Angulo, 1994; Rodríguez, 1995; Sanz, 1996; Alvira, 1996; Cruz, 1997; Sobrado y Ocampo, 1998). Disponemos incluso de estudios terciarios donde se indagan las realizaciones investigadoras españolas en diversos tópicos, utilizando metodologías cualitativas del tipo de la revisión integrativa-literaria (i.e. Fernández Cano, 1997), y cuantitativas (i.e. Fernández Cano y Bueno, 1998, 1999).

Además, la evaluación de programas educativos se encuentra en un momento de replanteamiento en el que las investigaciones de síntesis pueden arrojar cierta luz.

Fernández Cano y Bueno (1998 y 1999) en dos estudios terciarios sobre indicadores cientimétricos en investigación educativa distinguen tres grandes grupos: índice de productividad, de citación y relativos a contenido. A su vez los indicadores de productividad admiten cuatro subindicadores bien diferenciados: productividad diacrónica, de autores, institucional y de fuentes. Sin duda, es la productividad de autores el índice más estudiado y utilizado en la literatura cientimétrica, pudiéndose además diversificarse este índice en: distribución de estudios por autores tratando de verificar dos leyes cientimétricas comúnmente aceptadas (la ley de dispersión de Bradford, 1950, y la Ley de Lotka, 1927), clasificación de autores más productivos y multiautoría o grado de colaboración.

La adecuación del tópico que indagamos a patrones cientimétricos puede otorgar cierto carácter científico, por similitud, al tópico de la evaluación de programas y a los agentes que en él operan.

## **2. MÉTODO**

### **2.1. El problema. Objetivos del estudio**

Este estudio pretende realizar una síntesis de la investigación española acerca de la evaluación de programas educativos durante el periodo 1975/2000, mediante el estudio de los factores y leyes cientimétricas de productividad de los trabajos de investigación en este campo publicados como artículos en revistas científicas. Se trata de alcanzar el siguiente objetivo general:

- Ofrecer una visión panorámica acerca de la producción científica en este campo y periodo.

Y más específicamente en sus aspectos cientimétricos, los objetivos a alcanzar son:

- Ofrecer una visión diacrónica de la producción de las investigaciones sobre evaluación de programas educativos en el ámbito español, exponiendo su evolución.
- Denotar la productividad de los autores en cuantía y tipo.

- Indagar la filiación institucional de los autores.
- Señalar las fuentes más utilizadas para la publicación de estos trabajos científicos.
- Caracterizar la investigación educativa en este campo y periodo, según normas cientimétricas, comprobando si se verifican las leyes más relevantes a partir de los indicadores básicos de productividad.

## 2.2. Diseño

Este estudio encaja perfectamente en la definición que ofrece Colás (1998) de los estudios *descriptivos*, que tienen como principal objeto describir sistemáticamente hechos y características de una población o área de interés de forma objetiva y comprobable.

Además resultaría un estudio *analítico documental y de desarrollo*, puesto que se pretende analizar el fenómeno de la investigación española sobre evaluación de programas en algunos de sus componentes básicos, como son en este caso sus aspectos cientimétricos de productividad; y por otro lado, se pretenden buscar patrones de desarrollo a lo largo del periodo 1975/2000.

A la vez, posee un *carácter de síntesis* o estudio secundario, puesto que el contenido de estudio es la producción científica primaria, expresada por artículos publicados en revistas especializadas, acerca del tópico de interés y tratadas mediante una *metodología cientimétrica*.

Realizamos una aproximación no normativa. No pretendiendo el estudio de programas estándares, sino de sus informes de evaluación atendiendo a aspectos cientimétricos.

En cuanto a la validez del diseño para generalizar los resultados de esta revisión a toda la producción investigativa sobre evaluación de programas, o *validez de la población* (Hernández Pina, 1998), realizamos este proceso de inferencia con las precauciones derivadas que nuestra muestra no fuese lo suficientemente significativa de la población que pretende describir.

## 2.3. Variables

En los estudios descriptivos se utilizan indicadores expresados, por lo general, mediante estadísticos como frecuencias, porcentajes y medias aritméticas (Briones, 1991). Adaptados en estudios de tipo cientimétrico (Fernández Cano y Bueno, 1999) por los indicadores específicos referenciados a continuación y en los siguientes aspectos:

1. Productividad diacrónica: Número total de estudios seleccionados en cada año desde 1975 hasta el 2000.
2. Productividad personal, que incorpora los siguientes sub-indicadores:
  - Número de trabajos en solitario y número de trabajos en coautoría de cada uno de los autores seleccionados.

- Clasificación de productividad o listado de los autores con más publicaciones.
  - Número de autores promedio o índice de coautoría.
3. Productividad institucional: Número de estudios y porcentaje según la filiación institucional de sus autores.
  4. Productividad de la fuente. Haciendo referencia a la fuente en que son publicados cada uno de los artículos seleccionados se realiza una distribución, su frecuencia y el porcentaje del total de publicación.

## 2.4. Instrumentos

El instrumento para la recogida de datos se basaba en una hoja de registro informatizada mediante el programa Excel, en la que se registraba datos referidos a: autor, fecha, institución y fuente de cada publicación.

Este instrumento ha sometido a una dinámica de modificación y mejora durante el proceso de registro para adaptarlo mejor a los requerimientos de los distintos datos. Así, se han ido homogeneizando los campos de registro para que los datos de todos los artículos revisados pudieran quedar convenientemente registrados. Al ser éste un instrumento informatizado desde su propia concepción se han flexibilizado tanto las opciones de cambio, como facilitado las anotaciones y el posterior tratamiento en las hojas de análisis de datos y en general el tratamiento de la información.

La validez del instrumento puede quedar determinada, en alguno de los aspectos que determinan Colás y Buendía (1998), en dos momentos de la investigación: En cuanto a su contenido, durante el diseño se atendió a cuestiones relevantes expresadas en estudios afines y a la revisión de expertos. Y en su utilización, por el registro de múltiples observadores/registradores expertos que lo han ido optimizando a lo largo del proceso de registro.

En cuanto a su fiabilidad queda patente su consistencia al tratarse de un instrumento descriptivo de aspectos bibliométricos cuantitativos de valor único. Corroborado por valores próximos al 100% siempre obtenidos en triangulación de revisores, y con posterior obtención de consenso sobre datos discrepantes.

## 2.5. Muestra

La población a la que se pretende generalizar será la investigación educativa española relativa a la evaluación de programas. La muestra es intencional hasta pretender llegar al estudio causal, o sea, se trata de localizar para integrar todos los estudios empíricos al respecto.

La principal amenaza en esta revisión la podemos encontrar en la búsqueda y selección de la población y muestra. Es decir, en los posibles sesgos en la selección de los estudios primarios a revisar. Se intentó controlar realizando una selección a tres niveles o en tres fases.

En primer lugar por una búsqueda informatizada en la base de datos del I.S.O.C./C.S.I.C utilizando todas las combinaciones posibles de las palabras clave «evaluación + programas + educativo + españoles + investigación», complementándola por una búsqueda artesanal de las referencias cruzadas.

TABLA 1  
TÉRMINOS DE BÚSQUEDA Y REGISTROS ENCONTRADOS EN LA BASE DE  
DATOS ISOC / CSIC

Término de búsqueda	Registros encontrados
<i>Investigación</i>	9492
<i>Evaluación</i>	2458
<i>Programa</i>	3334
<i>Educación</i>	12199
<i>Español</i>	247
<i>Evaluación + programas</i>	160
<i>Programas + educación</i>	421
<i>Programas + educación + evaluación</i>	18

En segundo lugar, basándonos en la revisión de la literatura sobre evaluación de programas se establecieron unos criterios generales para la selección de los trabajos, que fueron los siguientes:

- Los programas debían ser programas que trataran cualquier aspecto educativo: desarrollo curricular, educación para la salud, para la convivencia, etc.; tanto en el ámbito escolar, familiar como social.
- Los programas o las evaluaciones debían realizarse en el ámbito español. Y si se referían a programas extranjeros, se habían aplicado y/o la publicación de su informe debía realizarse en revistas españolas.
- Que se tratase de evaluaciones reales realizadas a programas implantados totalmente, o estuviesen en alguna de sus fases.
- Las evaluaciones debían tener unos criterios mínimos de «investigación educativa», en la que se estableciesen unos objetivos o fines, metodología, instrumentos, se realizase una recogida y análisis de datos, que permitiesen emitir unas conclusiones y tomar decisiones acerca del programa. Desestimando los artículos de opinión, teóricos o con argumentaciones no fundamentadas.

Y por último, como las revisiones nunca son asépticas (Álvarez Rojo y Hernández, 1997) sino que implican una toma de decisiones por parte del revisor a la hora de seleccionar los trabajos a revisar, corriendo el riesgo de omitir alguno realmente significativo. Entonces, se utilizó un mecanismo de fiabilidad usual en los estudios de síntesis, en el que cada lectura era revisada detenidamente por dos recolectores/lec-

tores que decidían y resolvían conjuntamente las discrepancias sobre su selección definitiva.

Los problemas más habituales se dieron tras la búsqueda automatizada y la selección en el proceso de recuperación y nueva búsqueda manual de referencias, por cuestiones como:

- En algunas referencias de la base ISOC ciertos datos, como año de publicación o número de publicación, no se correspondían con los datos reales del trabajo.
- Nuestra atención se centraba en artículos de investigación sobre programas educativos españoles. El que algunas referencias se encontrasen en lenguas autonómicas, requirieron un esfuerzo para su traducción.
- Aunque todos los trabajos hacían referencia al mismo tópico o asunto, la dispersión de la información es evidente, teniendo que recurrir al préstamo interbibliotecario o al apoyo externo para recuperar referencias.

En total se seleccionaron 166 artículos de investigación, de los que se pudieron localizar, recuperar y revisar 138. Seis artículos fueron localizados pero tras una revisión preliminar desestimados por no realizar las citaciones y referencias bibliográficas completas o por ser una mera compilación de estudios y no una evaluación específica, y 22 artículos no pudieron ser localizados, en su mayor parte por la mortalidad de las fuentes en las que fueron publicados.

Asumimos además, el supuesto conceptual que asimila la investigación realizada con su presentación como artículo en revista científica. Este isomorfismo no es descabellado, ya que el artículo es el vehículo habitual por excelencia para difundir los informes de una investigación.

### **3. RESULTADOS**

#### **3.1. Productividad diacrónica**

Entendida como el número de artículos por cada año del periodo estudiado y en los periodos establecidos en el apartado de indicadores y variables del estudio. De acuerdo con los artículos revisados en este estudio obtenemos los siguientes totales y porcentajes de publicación, contenidos en la tabla 2.

Podemos apreciar un notable incremento a partir de 1986 en 40 artículos en términos absolutos. La producción se estabiliza durante dos periodos con valores similares, para descender en el 1996/2000. Esta tendencia podría tener una explicación, además del descenso en la producción, en la falta de actualización de las fuentes en que se han obtenido los datos para este estudio. Así, muchos de los artículos publicados durante este último año no estaban todavía registrados en las bases de datos durante el primer trimestre en que se realizó la revisión, por lo que se han debido localizar por medio de referencias cruzadas manuales.

Acumulando los porcentajes de publicación en los periodos diacrónicos podemos comprobar que en los tres primeros obtenemos tan sólo el 38,4 % de la producción y en

TABLA 2  
PRODUCTIVIDAD DIACRÓNICA POR PERIODOS

PERIODO DE PUBLICACIÓN	Nº ARTÍCULOS	%	% ACUMULADO
1975 / 1980	3	2.17	2.17
1981 / 1985	5	3.62	5.79
1986 / 1990	45	32.61	38.4
1991 / 1995	49	35.51	73.91
1996 / 2000	36	26.09	100
<b>Nº TOTAL</b>	<b>138</b>	<b>100</b>	<b>100</b>

los dos siguientes un 61,6 %. Es decir, es éste un tópico de interés en los dos últimos periodos temporales, como se aprecia en la figura 1.

Claramente en la productividad de trabajos acerca de evaluación de programas educativos se ha alcanzado el modelo de curva logística, como superación de la curva de incremento exponencial. Es decir la producción ha dejado de crecer exponencialmente para estabilizarse. Esto es un patrón de madurez científica (Price, 1986) propio de sistemas consolidados, pues el crecimiento continuo llevaría al absurdo.

Atendiendo a estos datos, podemos afirmar que la evaluación de programas educativos es un tópico de investigación vigente durante las dos últimas décadas en nuestro país y parece que todavía un frente activo, ya que no sólo habría que tener en cuenta el número de artículos publicados, sino la profundidad de éstos.

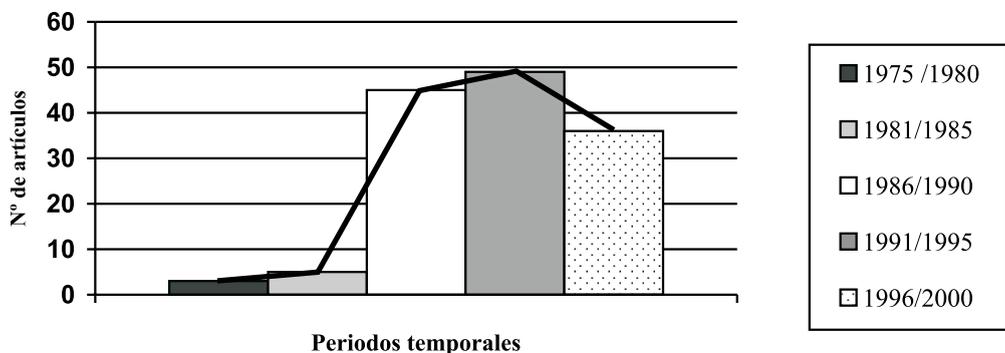


Figura 1  
Histograma de productividad diacrónica y línea de tendencia.

### 3.2. Productividad de autores

Con estos indicadores se pretende describir el número de trabajos realizados por los distintos autores sobre el tema tratado, para poder ordenarlos de acuerdo a su productividad.

TABLA 3  
CLASIFICACIÓN DE AUTORES POR SU PRODUCTIVIDAD

ORDEN	AUTORES	Nº TRABAJOS
1	Gargallo, B.	7
2	Trianes -Torres, M.V.	4
3	Babio, M.	3
4	Echeitia, G.	3
5	Galán, M.	3
6	Martín, E.	3
7	Muñoz-Sánchez, A.	3
8	20 Autores con...	2 (*)
9	231 Autores con...	1 (*)

(\*) Los autores ocasionales, con dos o un solo trabajo, que ascienden a 251 pueden consultarse en (Expósito, 2000).

Atendiendo al tamaño de esta distribución se puede comprobar que el mayor número de autores sólo tienen uno/dos trabajos; o sea, son autores ocasionales. También es importante describir la relación entre los trabajos que se realizan en solitario y los que se realizan mediante la colaboración de varios autores, coautoría. Resultando 59 trabajos en solitario y 79 en coautoría de los 138 revisados. Es decir, un 57%, la mayoría, frente a un 43% de los trabajos de investigación se realizan en coautoría.

El índice de coautoría o número de autores promedio, resulta de realizar el cociente entre el número total de autores, tomados uno a uno, por el número total de artículos revisados, en este caso 138. Obtenemos un valor de 1,9 autores/trabajo, no resultando un valor demasiado alto, que indique que el campo de la evaluación de programas se venga abordando con una metodología colaborativa; más bien, parece ser un trabajo de dos colaboradores, que bien pudieran ser, conjeturamos, doctorando y director de tesis propios de ese artículo.

Se puede comprobar que la distribución de los autores según las aportaciones autor/trabajos, sigue los siguientes valores (distribución observada o patrón empírico) en comparación con lo indicado en la Ley de Lotka (distribución esperada o patrón teórico).

Para comprobar los ajustes entre distribuciones suelen utilizarse estadísticos que las correlacionen: En este caso se puede comprobar una correlación positiva entre ambas distribuciones con un índice de correlación ordinal de  $J = + 0,99$  y un coeficiente  $r = 0,98^2$ .

TABLA 4  
ADECUACIÓN DE LA PRODUCCIÓN PERSONAL A LA LEY DE LOTKA

Nº de contribuciones / autor	Nº de autores (Esperados)	Nº de autores (Observados)	%
1	231	231	89.8%
2	57.75	20	7.7%
3	25.6	5	1.9%
4	14.43	1	0.3%
7	4.71	1	0.3%

En términos absolutos se puede comprobar que el 89.9%, la mayoría, de los autores realizan tan sólo una aportación o trabajo. Atendiendo a la clasificación de productores propuesta por (Crane, 1969), el 89,9% serían considerados productores ocasionales o transeúntes, tan solo siete autores realizan más de tres contribuciones o pueden ser considerados medianos o moderados productores. No existiendo ningún autor que realice más de diez aportaciones, para poder ser considerado un gran productor. Es decir, podemos afirmar, que son unos pocos autores los que más aportaciones realizan en la investigación de este tópico o la producción se encuentra en manos de unos pocos.

### 3.3. Productividad institucional

Atendiendo a la afiliación institucional de los autores encontramos los siguientes datos y parámetros de producción en los 137 artículos en los que se han podido recopilar este campo.

De forma absolutamente clara, son las universidades las que soportan la mayoría de la investigación sobre evaluación de programas en un alto porcentaje 65,69%, siendo también destacable el escaso porcentaje obtenido por las diversas administraciones tanto a nivel central, como autonómico y local.

Atendiendo a estos porcentajes podríamos distinguir Centros de excelencia de productores, al menos cuantitativamente, en la evaluación de programas que serían la Universidad de Valencia, Complutense de Madrid, Central de Barcelona y la Universidad de Granada.

Por otro lado, se puede apreciar una escasa aportación a la producción por parte de equipos empresariales y editoriales, así como en proyectos de colaboración internacional.

### 3.4. Productividad por fuentes de publicación

La fuente de publicación se refiere a la revista en la que son publicados los artículos revisados para este estudio. Señalando las diez más productivas en este campo.

TABLA 5  
PRODUCTIVIDAD INSTITUCIONAL EN NÚMERO DE TRABAJOS Y PORCENTAJE

Instituciones productoras		Nº	%	%
UNIVERSIDAD	Valencia	13	9,49	65,69
	Complutense (Madrid)	11	8,03	
	U.N.E.D.	9	6,57	
	Central (Barcelona)	8	5,84	
	Granada	7	5,11	
	Málaga	6	4,38	
	Oviedo / País Vasco / Salamanca / Sevilla	(4 c.u.) 16	11,68	
	Autónoma (Madrid) / La Laguna (Tenerife)	(3 c.u.) 6	4,38	
	Almería / Murcia / Orense	(2 c.u.) 6	4,38	
	Alcalá Henares / Castellón / Castilla La Mancha / Córdoba / Girona / León / Navarra / Vigo	(1 c.u.) 8	5,84	
ADMINISTRACIONES	M.E.C.	3	2,19	10,22
	Autonómicas	7	5,11	
	Locales	4	2,92	
OTRAS	Centros educativos no universitarios	9	6,57	19,71
	Equipos de apoyo educativo	4	2,92	
	Institutos y equipos oficiales	8	5,84	
	Equipos biosanitarios	2	1,46	
	Equipos editoriales / empresariales.	4	2,92	
COLABORACIÓN INTERNACIONAL		6	4,38	4,38
TOTAL		137	100 %	100%

La mayor publicación de artículos de investigación acerca de la evaluación de programas educativos corresponde a la *Revista de Investigación Educativa (R.I.E.)*, que con un porcentaje del 22,46 del total supera en más de diecisiete puntos a la siguiente, que es la *Revista Española de Pedagogía*, que posee un 5,07%. Para igualar su producción es necesario sumar las de las seis revistas siguientes en la clasificación.

Existe un conglomerado de publicaciones; *Bordón*, *Revista de Ciencias de la Educación* y *Revista de Educación* que con 6 artículos obtienen un 4,32 de la producción. Y otro formado por *Adicciones*, *Infancia y Aprendizaje* y *Tarbiya* que con 4 publicaciones obtienen un porcentaje de 2,88.

Hay que tener en cuenta que algunas de estas revistas aunque podían ofrecer en su producción artículos referentes a la evaluación de programas durante el periodo estudiado, no cumplían las condiciones establecidas para la selección.

TABLA 6  
RANKING EN PRODUCTIVIDAD DE LAS FUENTES (REVISTAS)

Nombre publicación	Nº trabajos	%
1ª R.I.E.	31	22,46
2ª Revista Española de Pedagogía	7	5,07
3ª Bordón	6	4,32
4ª Revista de Ciencias de la Educación	6	4,32
5ª Revista de Educación	6	4,32
6ª Adicciones	4	2,88
7ª Infancia y aprendizaje.	4	2,88
8ª Tarbiya	4	2,88
9ª Con dos artículos (*)	12 X2= 24	17,39
10ª Con un artículo (*)	46 X1= 46	33,33

La distribución de fuentes con dos o un solo artículo seleccionado puede consultarse en (Expósito, 2000).

Es importante señalar que un total de 58 estudios aparecen publicados en revistas que tan sólo poseen uno o dos trabajos de este tipo en el periodo estudiado; por lo que, la mayor cuantía de la producción en términos absolutos se realiza en fuentes de publicación que podríamos calificar de ocasionales en cuanto a publicación referente a este tópico.

Existen dos grupos de 12 y 46 publicaciones que editando dos y sólo un solo trabajo respectivamente, que cumplan las condiciones de selección en esta revisión; obteniendo en su conjunto el 17,3% y 33,3% del total de la producción (Expósito, 2000).

TABLA 7  
ADECUACIÓN DE LA PRODUCCIÓN POR FUENTES A LA LEY DE LOTKA

Nº de contribuciones / fuente	Nº de autores (Esperados)	Nº de autores (Observados)	%
1	46	46	61,33
2	11,5	24	32
4	2,87	3	4
6	1,27	3	4
7	0,93	1	1,33
11	0,04	1	1,33

### 3.5. Ajuste a leyes cientimétricas

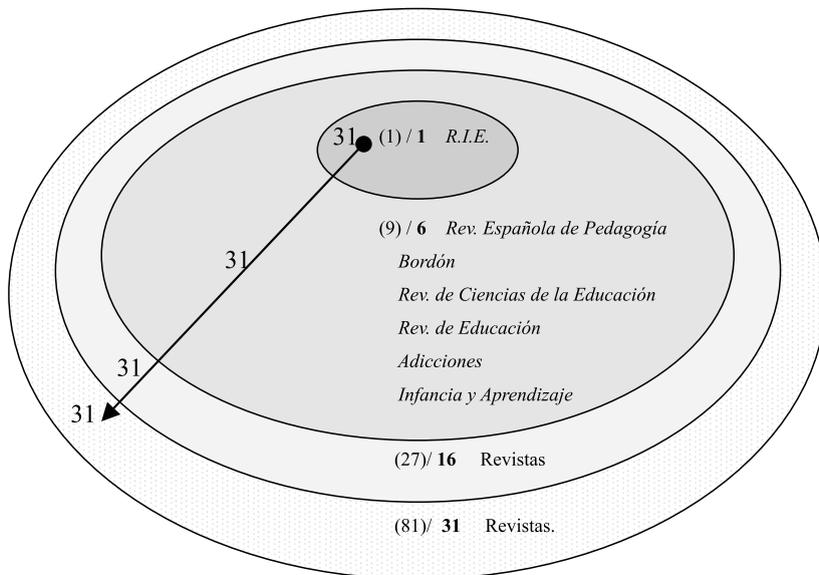
Para comprobar los ajustes entre distribuciones suelen utilizarse estadísticos que las correlacionen: En este caso se puede comprobar una correlación positiva entre ambas distribuciones con un índice de correlación ordinal de  $J = + 0,58$  y un coeficiente  $r = 0,33^2$ .

Podemos comprobar que la mayoría de la producción también está en manos de unas pocas fuentes, que podemos agrupar en áreas de producción o zonas de Bradford. Así, la distribución del número de trabajos por el número de revistas que publican esos trabajos es la registrada como *observada*, y su adecuación a la distribución según la Ley de Bradford o *esperada* es la siguiente:

TABLA 8  
ADECUACIÓN DE LA PRODUCCIÓN A LA DISTRIBUCIÓN DE BRADFORD

Zonas	1ª	2ª	3ª	4ª
Producción	31	31	31	31
Nº de revistas (Esperado)	1	9	27	81
Nº de revistas (Observado)	<b>1</b>	<b>6</b>	<b>16</b>	<b>31</b>
FUENTES	<b>R.I.E.</b>	Revista Española de Pedagogía Bordón Revista de Ciencias de la Educación Revista de Educación Adicciones Infancia y Aprendizaje	*	*

(\*) Estas distribuciones nominales pueden consultarse en (Expósito, 2000).



Distribución (teórica)/experimental.

Figura 2  
Representación de las fuentes y su producción en las zonas Bradford.

Para mantener una cantidad de producción acerca del tópico investigado similar a la de la revista R.I.E. es necesario unir la producción de las seis revistas siguientes en la clasificación de producción. A su vez para mantener un número similar habría que asociar la producción de las 16 siguientes publicaciones. Y en un nuevo conglomerado de 31 de las que están a continuación.

Así se forman conglomerados de publicaciones o *zonas de Bradford* alrededor de una revista nuclear, de mayor producción, según la cantidad de artículos publicados acerca del tópico de investigación. En este caso podemos comprobar que la *Revista de Investigación Educativa* se convierte en un referente en cuanto a la investigación en este campo.

#### 4. CONCLUSIONES

En cuanto a la productividad diacrónica de la investigación, representada ésta por los artículos revisados, en evaluación de programas educativos en España durante el periodo 1975-2000, podemos considerar:

La producción muy baja entre 1975 y 1985 aumenta considerablemente desde 1986 hasta 1995, para decaer en el último periodo 1996/2000. Atendiendo a la curva de tendencia se puede comprobar que se ha dejado un momento de crecimiento de la producción exponencial, par estabilizarse.

En estos momentos más que esperar un aumento cuantitativo de la producción, parece lógico y deseable esperar su mejora cualitativa en aspectos fundamentales como la difusión e impacto de los hallazgos que la investigación produce. Esta es sin duda, como afirma Fernández Cano (2000 a) una de las cuestiones pendientes en la investigación educativa en nuestro país.

Hay que considerar que la cantidad de la producción registrada en el último periodo puede verse sesgada por el término del proceso de registro en el primer cuatrimestre del año 2000. Ya que la base de datos consultada es lenta para ir incluyendo informaciones actualizadas y que las revistas no son puntuales en su publicación, puede darse el caso de trabajos que, aunque realizados no estén registrados, y no hallan podido ser detectados para su selección.

La producción científica en este campo podemos calificarla de dispersa en cuanto a la autoría, como demuestra su ajuste a la Ley de Lotka. Son escasos los autores que pueden considerarse autores relevantes en cuanto a su constancia en el tópico, tan sólo Gargallo y Trianes-Torres poseen más de tres publicaciones. Hay un grupo de medianos productores formado por autores como Babio, Echeitia, Galán, Martín y Muñoz Sánchez, pero la mayoría son autores ocasionales con uno o a lo sumo dos trabajos.

Esta ocasionalidad en la producción en un campo científico disciplinar puede denotar por un lado el interés por parte de múltiples autores, aunque sea de forma ocasional; pero por otro, la debilidad de planteamientos estables en cuanto a la autoría por expertos en el tema.

Los trabajos realizados obtienen un índice de autoría promedio de 1,9 autores siendo éste un índice habitual en la investigación educativa en nuestro país (Fernández Cano y Bueno, 1998; 1999) aunque lejos de valores de autoría más altos propios de las ciencias «duras» (López Piñero y Terrada, 1992). Habría que comenzar a plantearse

cuestiones acerca de la conveniencia de dejar en algunas ocasiones el trabajo de investigación excesivamente individualista, de típico corte pre-humanista; para establecer mecanismos en los que por medio de la colaboración, más típico en la producción científica de otras disciplinas, permitiesen la realización de proyectos más ambiciosos. Se echa en falta pues esos «colegios invisibles» por productividad, mencionados por Crane (1969), que superen el mal endémico del aislamiento o el solipsismo tan propio de los investigadores educativos españoles.

La productividad institucional es básicamente universitaria (66%) frente a la no universitaria (34%). Y en un escalonamiento de las instituciones universitarias más productivas en cuanto al tópico de revisión serían: Valencia, Madrid-Complutense, Universidad Nacional de Educación a Distancia (U.N.E.D.), Barcelona-Central y Granada.

En cuanto a las fuentes de los trabajos publicados podemos comprobar que la mayoría de la producción también se publica en un núcleo relativamente pequeño de publicaciones. En concreto, en nuestro país existe un referente general en investigación educativa, y también específicamente en evaluación de programas, que es la revista R.I.E. Esta publicación estaría en la zona nuclear de la distribución propuesta por Bradford. Existiendo también un grupo de publicaciones con cierto interés por el tema como Bordón, Revista de Ciencias de la Educación, Revista de Educación e Infancia y Aprendizaje; y alguna muy especializada en evaluaciones de programas específicos como Adicciones, que ocuparían una zona concéntrica exterior en cuanto a su producción.

Sin embargo, la mayoría de las fuentes podrían calificarse de fuentes ocasionales, con elevada mortalidad o escasa difusión.

En conclusión podemos comprobar que la evaluación de programas educativos en nuestro país es un tópico vigente, en el que su productividad ha transformado su crecimiento de exponencial a logístico, proceso que algunos autores identifican con etapas de asentamiento y mejora (Fernández Cano, 2000b). La mayoría de los trabajos quedan en manos de un número relativamente limitado de autores y fuentes, siendo el resto productores/publicadores ocasionales.

## BIBLIOGRAFÍA

- Álvarez Rojo, V. y Hernández Fernández, J. (1997). El modelo de intervención por programas. Aportaciones para una revisión. *Revista de Investigación Educativa* (16), 2, 79-123.
- Alvira, F. (1996). *Metodología de la evaluación de programas*. Madrid: CIS.
- Arruabarrena, M. y Paul, J. (1994). Los programas de tratamiento de familias con problemas de maltrato y abandono infantil: Descripción y evaluación. *Infancia y Aprendizaje*, 71, 159-178.
- Bartolomé Pina, M. (1990). Evaluación y optimización de los diseños de intervención. *Revista de Investigación Educativa*, 8 (16), 39-59.
- Bradford, S.C. (1950). *Documentation*. Washington, D.C.: Public Affair Press.
- Briones, G. (1991). *Evaluación de programas sociales*. México: Trillas.
- Colás, P. (1998). Los métodos descriptivos. En *Investigación Educativa*, P. Colás y L. Buendía (3ª Edición), (pp. 177-200). Sevilla: Alfar.
- Colás, P. y Buendía, L. (1998). *Investigación Educativa*. (3ª Edición). Sevilla: Alfar.

- Colás, P. y Rebollo, M.A. (1993). *Evaluación de Programas. Una guía práctica*. Sevilla: Kronos.
- Crane, D. (1969). Social structure in a group of scientist: A test of 'invisible college' hipotesis. *American Sociological Review*, 34, 335-352.
- Cruz, J.M. (1997). *Evaluación de programas de educación para la carrera*. Sevilla: Kronos.
- De la Orden, A. y Martínez de Toda, M.J. (1992). Metaevaluación educativa. *Bordón*, 43(4), 517-527.
- Expósito López, J. (2000). *La investigación sobre evaluación de programas educativos en España 1975-2000. Aspectos cuantitativos*. Granada: Autor.
- Fernández Cano, A. (1995). *Métodos para evaluar la investigación*. Madrid: Síntesis.
- Fernández Cano, A. (1997). Evaluación de la investigación educativa española: Una revisión integrativa de realizaciones en 25 años. *Revista Española de Pedagogía*, 207, 277-301.
- Fernández Cano, A. y Bueno, A. (2000). Metodología de la evaluación de programas educativos. Documento interno (programa de la asignatura). Granada: Facultad de Ciencias de la Educación.
- Fernández Cano, A. y Bueno, A. (1999). Synthesizing scientometric patterns in Spanish educational research. *Scientometrics*, 46 (2), 349-367.
- Gelpi, E. (1987). *Técnicas de evaluación y seguimiento de programas de formación profesional*. Madrid: Largo Caballero.
- Hernández Pina, F. (1998). Diseños de investigación experimental. En *Métodos de investigación en psicopedagogía*, L. Buendía, P. Colás, y F. Hernández Pina, (pp. 2-60). Madrid: McGraw Hill.
- López Piñero, J.M. y Terrada, M.L. (1992). Los indicadores bibliométricos y la evaluación de la actividad médico-científica. Uso y abuso de la bibliometría. *Medicina Clínica*, 98, 64-68.
- Lotka, A.F. (1927). The frequency distribution of scientific productivity. *Journal of the Washintong Academy of Science*, 16, (June, 19), 317-323.
- Medina, A. y Angulo, L.M. (Coord.) (1994). *Evaluación de programas, centros y profesores*. Madrid: Universitas.
- Ortega y Gasset, J. (1929/1983). *La rebelión de las masas*. Barcelona: Orbis
- Price, D.J. de S. (1986). *Little science, Big Science... and beyond*, (edición aumentada). Nueva York: Columbia University Press.
- Repetto, E. (1987). Evaluación de programas en Orientación. En V. Álvarez (Ed.) *Metodología de la Orientación Educativa*. Sevilla: Alfar.
- Repetto, E. (1993). Modelos actuales de programas para el desarrollo de la carrera. Actas de I Jornadas de la AEOEP. Valencia: AEOEP.
- Repetto, E. (1995). La Orientación como intervención psicopedagógica. Avances relevantes en la investigación de algunos modelos. *Revista de Investigación Educativa*, 26, 233-252.
- Rodríguez, M.L. (1993). Programas de orientación para la transición escuela-trabajo. Actas de I Jornadas de la AEOEP. Valencia: AEOEP.
- Rodríguez, M.L. (1995). *Orientación e intervención psicopedagógica*. Barcelona: CEAC.
- Salvador, A. y Periró, J.M. (1986). *La madurez vocacional. Evaluación, desarrollo y programas de intervención*. Madrid: Alhambra.

- Sanz, R. (1996). *Evaluación de programas en orientación educativa*. Madrid: Pirámide.
- Scriven, M.S. (1974). Standards for the evaluation of educational programs and products. En *Evaluating educational programs and products*, G. Boric (Ed.). Englewood Cliffs: N.J.: Educational Technology Publications.
- Shadish, W.R.; Cook, TH.D. y Leviton, L. (1991). *Foundations of program evaluation. Theories of practice*. Newbury Park, CA: Sage.
- Sobrado, L. y Ocampo, C. (1998). *Evaluación psicopedagógica y orientación educativa*. Barcelona: Estel.
- Stufflebean, D.L. (1978). Metaevaluation. An overview. *Evaluation and the Health Professions*, 1 (2), 17-43.
- Stufflebean, D.L. (1981). Metaevaluation: Concepts, standards and uses. En *Educational evaluation methodology: The state of the art*. Baltimore, R.A. Berk (Ed.). Nueva York: Johns Hopkins University Press.
- Velaz de Medrano, C.; Blanco, A. y otros (1996). *Evaluación de programas y Centros educativos. Diez años de investigación*. Madrid: C.I.D.E.
- Worthen, B.R. y Sanders, J.R. (1987). *Educational evaluation. Alternative approaches and practical guidelines*. Londres: Longman.

Fecha de recepción: 15 de marzo de 2001.

Fecha de aceptación: 11 de enero de 2002.