International Journal of Parents in Education
2008, Vol..2, No. 1, 1-12

ISSN: 1973 - 3518

School-parents relationships: a bibliometric study on 40 years of scientific publications.

Stefano Castelli
Università degli studi Milano - Bicocca
Milano, Italy

Alessandro Pepe
Università degli studi Milano – Bicocca
Milano, Italy

The present paper explores the amount of research activity estimated through the growth of scientific literature in the field of parents' participation in their childrens' schooling: it is a so-called bibliometric analysis. Assessing the "size of science" is a complex task, but bibliometric techniques have been shown to be useful in the development of scientific research activity to address emerging concerns such as the rate of development of a certain discipline and of some specific areas of scientific interest (Katz & Hicks, 1997).
Data for analysis were collected from several relevant queries restricted to articles accessible through Cambridge Scientific Abstract (CSA) on-line research platform. Nearly 11,000 bibliographic records were retrieved, with publication dates ranging from 1966 to 2005.
In this study researchers attempted to answer two main research questions: which is the rate of literature growth in the field of parent participation (and the expected rate for the forthcoming years)? Which are the different patterns of literature production through different research fields (Pedagogy, Psychology, Sociology, etc.)? Our results reveal a strong growth rate in the number of articles published, with rather significant differences in the various disciplines taken into consideration.

Introduction

In scientific literature about school systems and children's learning achievements we find an ever-growing interest in the role played by parents in their relationships with the school environment. In the last twenty years, an enormous amount of scientific papers, books and reports have been published about the efforts enacted by politic authorities at all levels in order to include an increasing number of families in the education of their children.

But these are just qualitative and somehow subjective impressions. In order to give a sounder base to them, in this paper researchers decided to perform a bibliometric study of scientific literature about school-parents relationships.

Correspondence concerning this article should be addressed to Stefano Castelli, e-mail: Stefano.Castelli@unimib.it

From this effort, using quantitative evidences, we expected to answer a few relevant questions to the scholar in this field: when did the interest towards this subject started? Which is exactly the shape of this apparent growth? Which terms are most often used by authors when writing about school-parents interactions? Which are the disciplines where parents-school partnership is more extensively studied?

We used some of the most influential databases of scientific literature, i.e. those collected by Cambridge Scientific Abstracts (CSA), retrieving some 11.000 results starting from 1966 to 2005, with a representative selection of strings for the queries. The decision of using five-year intervals is due to the fact that CSA allows to make queries on yearly basis only from 1987 onwards. We are still working on a reasonable way to interpolate data, but in this paper we decided to use only comparable data based on five-year intervals.

Data have been analyzed and evaluated for what regards their distribution in time, and their distribution in different databases (psychologically-oriented databases vs. pedagogical and sociological ones).

The paper is divided in three parts. The first part is methodological in character and, besides some words devoted to illustrate possible models of growth in scientific literature, will discuss the many limitations of a bibliometric approach, along with its advantages. In part two, an overall description of data will be given, with a comparison between "psychological" vs. "non psychological databases". In part three some specific considerations will be offered to everybody who has some interest in the concept of parental involvement.

Bibliometric approach and models of growth

The term meta-analysis refers to a set of statistic procedures aimed at synthesizing results from multiple studies with quantitative data. But a meta-analytic study to assess the impact of parents' participation on school processes is nowadays redundant, since many other studies have already shown the relevance of conjoint activity of parents and teachers for what regards children's achievements (Fan & Chen 2001). The present study separates a little form the meta-analytic tradition, to embrace a bibliometric approach.

Bibliometrics is a set of methods used to study or measure texts and information. Widely spread in the fields of hard sciences, only recently some scholars have begun to explore the possibility of developing bibliometric indicators for social sciences (e.g., Glanzel & Schoepflin, 1999; Hicks, 1999; Katz, 1999; Van Leeuwen, 2006).

Of course, we agree with Katz (1999) when writes: "Bibliometric indicators cannot capture all knowledge production in a society and inform us of its quality" (p.1). The main bias of bibliometric indicators is that they take mostly into account the results published in research journals. It is not possible to evaluate other – however important - contributions, i.e. publications in the form of books, best practices or research reports. At the same time, the use of a simple count of occurrences in a database is not able to capture the incidence of the results in the different fields of practical work.

Another aspect also worth of consideration is that a bibliometric study, for obvious reasons due to the characteristics of available databases, tends to be not representative of all publications. It is easy to find a predominance of publications in English language and in journals with a high impact factor. Thus, many original and interesting contributions to the development of a certain discipline remain excluded from the analysis.

Moreover, for what regards social sciences in particular, it is easy to witness a lower diffusion of journals and conference proceedings.

In spite of all these cultural and methodological biases, the attempts of measuring the size of science began already around Mid-Twenties (Fernandez-Cano, Torralbo & Vallejo 2004). It is important to stress the distinction between the two concepts of growth of size of science and of growth of knowledge. The former concept is easy to define and to measure, the latter presents a higher level of abstraction and complexity, and its evaluation is really difficult.

The most important model of growth of scientific production has been proposed by De Solla Price (1965), a pioneer in bibliometric research. The model postulates a strong relationship between the size of scientific production in a certain discipline and the number of years from its origin. De Solla Price considered the growth of scientific literature as similar to the biological growth of a population of bacteria. The growth of an organism tends to be a direct function of its size: the bigger it becomes, the faster it grows. Only in later years he introduced in his model the saturation hypothesis, thus theorizing a growth of knowledge in three major phases:

a) a preliminary phase, where increments are small;
b) a central phase of exponential growth of literature;
c) a final phase of stable and steady growth.

Even if the model had to be further elaborated (Fernandez-Cano, Torralbo & Vallejo 2004), the idea of considering the growth of science as a biological system is undoubtedly fascinating. More recently, Tague, Beheshti and Rees-Potter (1981) proposed a somewhat different approach, according to which three different curves can be used in modelling scientific production.

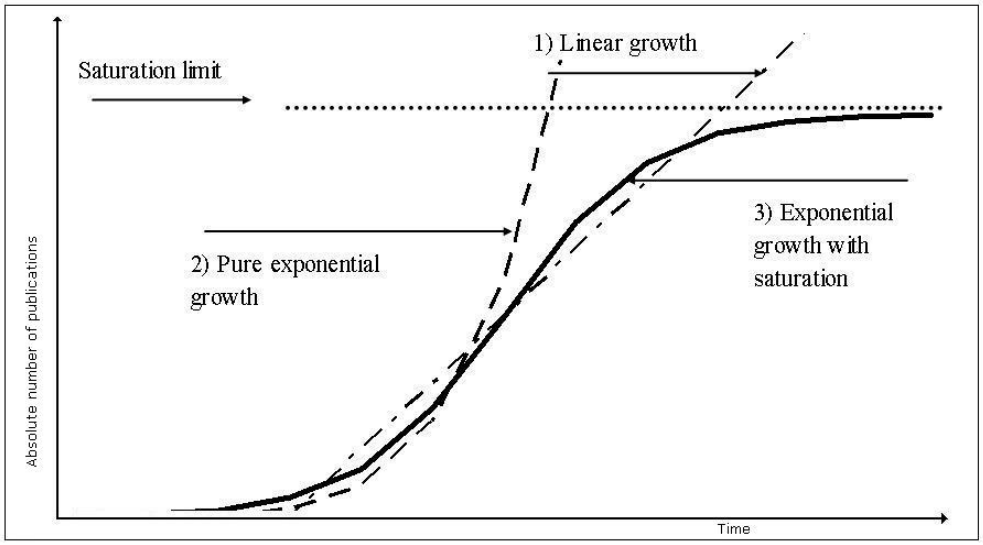


Figure 1. Three theoretical models for the growth of scientific production.

A linear model can be used to approximate growth in the short period. In an exponential model it is assumed that growth takes place without difficulties, and can be modeled through an exponential function. In a logistic model the function starts as an exponential curve but the growth slows down when a certain saturation level is reached. Figure 1 shows the three theoretical models for the growth of scientific production in a certain field of research.

Data Gathering and Methodology

Data for analyses have been collected in March 2007 through Cambridge Scientific Abstracts (CSA) research platform with a selection of the most important possible queries. A widely used way to conduct bibliometric studies is through Social Science Citation Index (SSCI), presently managed by Thomson Scientific (Gu, 2004). But for what regards social sciences, SSCI presents strong limitations of coverage (Van Leeuwen, 2006). Of course, no bibliographic database is free from the biases already mentioned before. But CSA seems to be less biased than other research platforms; however, it has to be stressed that our results are valid only within this set of databases.

CSA has a search interface which allows to search simultaneously many databases. In the present study, the following databases have been used: PsycINFO, Applied Social Sciences Index and Abstracts (ASSIA), CSA Social Service Abstracts (CSA-SSA), CSA Sociological Abstracts (CSA-SA), Educational Resources Information Center (ERIC). For a description of the databases, see Table 1 (Appendix 1).

As shown in Table 2, the databases comprise a total of about 4.6 million entries, published in nearly 5,700 journals. All databases host publications also in languages other than English.

A decision was made to divide the databases, according to their content, in two groups i.e. "psychological" (PsycINFO, and ASSIA) and "non-psychological" (Social Service Abstract, Sociological Abstracts, ERIC). For sure, this segmentation is somewhat problematic, since both groups show some overlap with respect to the journals covered. However, this seemed a reasonable way to put some order in a vast material. A deeper insight on the overlap could be the object of further research. Anyway, in order to avoid an overestimation of some results, all duplicate entries (i.e., entries appearing in more than one database) have been removed; this, obviously, entails the risk of an underestimation of the number of articles.

Three data matrices have been built; two matrices contained the number of occurrences retrieved by a query with a certain string, separately recorded for "psychological" and "non-psychological" databases. The third matrix, with the total amount of occurrences, was useful to obtain a comprehensive view, and to cross-check the data of the other two matrices.

To explore the original dataset, a selection of the most widely diffused strings in literature

Table 2.

A synthesis of characteristics of Cambridge Scientific Abstracts (CSA)

Database	*Size*	*Year Coverage*	*Number of Journals*	*English /Other Languages*
PsycINFO	2,271,000	1840-2005	2,097	yes/yes
ASSIA	375,000	1987-2005	500	yes/yes
CSA - SSA	107,000	1979-2005	1,300	yes/yes
CSA - SA	680,000	1963-2005	1,800	yes/yes
ERIC	1,172,000	1965-2005	n.a.	yes/yes

Note. Last available data from comparable queries between databases, December 2005

has been used. When in the search string there was no explicit mention of the word "school", we added it in the query through the Boolean operator AND, in order to refine the search excluding non relevant results.

The search strings used for the queries are listed in Table 3 (Appendix 1), together with the name in which the variable was coded.

Results

For what regards the productivity of the search strings, in Table 4 the absolute values of the results obtained by the different queries are summarized. From this table, we learn that the use of the strings parental involvement (par_inv, n=5,438) and family involvement (fam_inv, n=2,888) is a very fruitful research strategy, which allows to find a total of 8,326 publications (76,6%). On the contrary, other strategies prove by far less efficient.

Table 4.

Overall results grouped by search queries

Query	*Occurrences*	*Cumulative %*	*Query*	*Occurrences*	*Cumulative %*
Par_inv	5,438	50,0	Pt_com	114	1,0
Fam_inv	2,888	26,6	Hs_par	101	0,9
Par_par	635	5,8	Hs_con	83	0,8
Prs_inv	552	5,1	Hs_rel	80	0,7
Fam_par	394	3,6	Pt_rel	72	0,7
Prs_par	262	2,4	Pt_par	25	0,2
Hs_coll	228	2,1	**Total**	10.872	100,0

Table 5 shows the overall number of occurrences of all research strings divided in five-year intervals.

The total number of all publications from 1966 to 2005 is 10,872. Starting with only 121 publications in the first quinquennium (1966-1970), a number of 4,273 is reached in the last considered five-year period 2001-2005: the increase rate is 35 times in forty years.

For what regards distribution in time, there is a strong concentration in the last fifteen years: from 1991 to 2005, 8,457 occurrences were counted (i.e. 77.8% of the total number). The Nineties seem to be characterized by a strong increase in publications on parental involvement (par_inv).

The field is still in its "youth"; in fact, about 4,200 papers of the total (39.3%) have been published in the quinquennium 2000-2005. This fact could account for some methodological difficulties indicated in literature (Baker & Soden, 1998).

In order to evaluate if the observed values are compatible with some of the models exposed in the first part of this paper, we can start to plot absolute values in the whole dataset (Figure 2). Obviously, the growth is not linear, so we tested only for exponential and logistic curves. To evaluate the goodness of fit, we used R^2 (a standard measure provided by SPSS 14).

The results are very encouraging (R^2 > .90, in all our attempts to fit different logistic curves) with upper bounds – i.e. saturation limits – for the logistic curve ranging from 5,000 to 10,000 publications/quinquennium. In the case of no upper bound set, the results for all the three databases (psychological, non-psychological, total) range from .90 to .99, with R^2 identical both for the logistical and the exponential model. Of course, in future studies we will need to use techniques of cross-validation between models, and a higher number of data (i.e., points in the curve). But the evidence of an exponential growth (or, more likely, of the first phase of exponential growth in a logistic curve) is a strong element which stress the need for further investigations.

Table 5.
Overall number of occurrences grouped by five-years interval

Years	*Reference retrieved*	*Cumulative Value*	*Cumulative %*
1966-1970	121	121	1.1
1971-1975	245	366	3.4
1976-1980	470	836	7.7
1981-1985	585	1,421	13.1
1986-1990	994	2,415	22.2
1991-1995	1,707	4,122	37.9
1996-2000	2,477	6,599	60.7
2001-2005	4,273	10,872	100.0

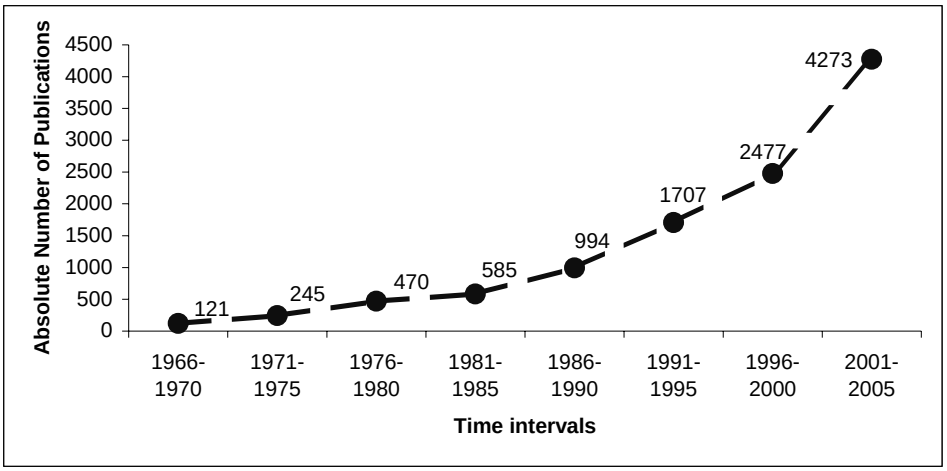


Figure 2. Absolute number of publications retrieved in each 5-year period.

Given this trend, it is not surprising that differences between the amount of publications retrieved in the global sample in a certain five-year period confronted with the amount found in the immediately previous period are always positive (see Figure 3). This indicates a continuous growth. However, things change dramatically if we split the databases, and consider separately the databases of psychological character and the ones dedicated to sociology and pedagogy. The results are presented in Table 6.

Table 6.
Overall results grouped by Psychological vs. Non-psychological database

	Psychological			*Non-psychological*		
Years	*Reference retrieved*	*Cumulative Value*	*Cumulative %*	*Reference retrieved*	*Cumulative Value*	*Cumulative %*
1966-1970	4	4	0.1	117	117	2.2
1971-1975	37	41	0.7	208	325	6.2
1976-1980	84	125	2.2	386	711	13.5
1981-1985	197	322	5.7	388	1.099	20.9
1986-1990	347	669	11.9	647	1.746	33.2
1991-1995	553	1,222	21.8	1,154	2,900	55.0
1996-2000	1,112	2,334	41.6	1,365	4,265	81.0
2001-2005	3,282	5,616	100.0	991	5,256	100.0

Note. "Psychological": PsychINFO, ASSIA;"Non-Psychological": CSA - SSA, CSA -SA, ERIC

The total amounts of findings are similar (5,616 vs. 5,256), but the distribution is very different. If we consider the last five-year period 2000-2005, there are 3,282 publications in psychological areas (59.4%), against only 991 (19.1%) in non-psychological databases. Taking into account previous periods, in this case we witness an inversion for what regards predominance of the two areas. In fact, until year 2000 publications in non-psychological areas constantly outnumbered psychological publications.

Cumulative percentages inform us that the distribution in psychological databases is very strongly concentrated in the years from 1991 to 2005 (nearly 80% of all publications).

Figure 4 plots the trends of growth divided according to the kind of databases (diamonds are psychological databases, squares the non-psychological ones); as in Figure 3, also here the differences between the amount of publications in a certain five-year period confronted with the amount found in the immediately previous period are plotted.

For what regards the psychological area, there is a constant trend of growth, with a minimum during the quinquennium 1971-1975 (+33 papers) and a maximum between 2001 and 2005 (+2,170). Non-psychological databases show a completely different trend: here the maximum is reached in 1991-1995 (+507), to slow down in 1996-2000 (+211), a tendency further stressed in the last period (-374). It is interesting to note a coincidence in time between the decrease in non-psychological databases and the strong increase of publications in psychological areas. We could be tempted to speak of a sort of "migration" from one disciplinary area to the other, but presently data don't allow us to do so; in fact, the production in non-psychological areas shows other periods of slowing down (for instance, in 1981-1985 there is a growth of only +2 papers).

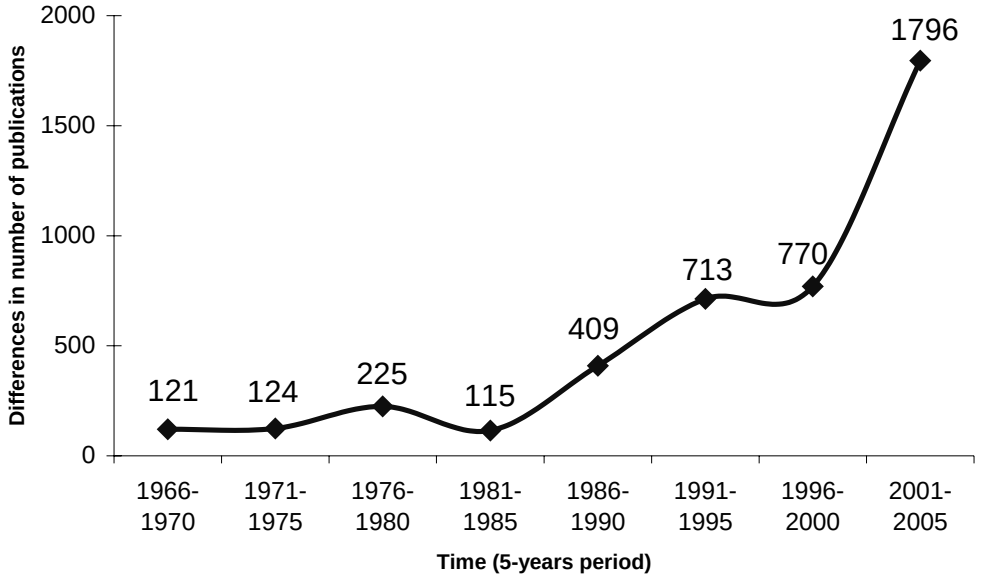
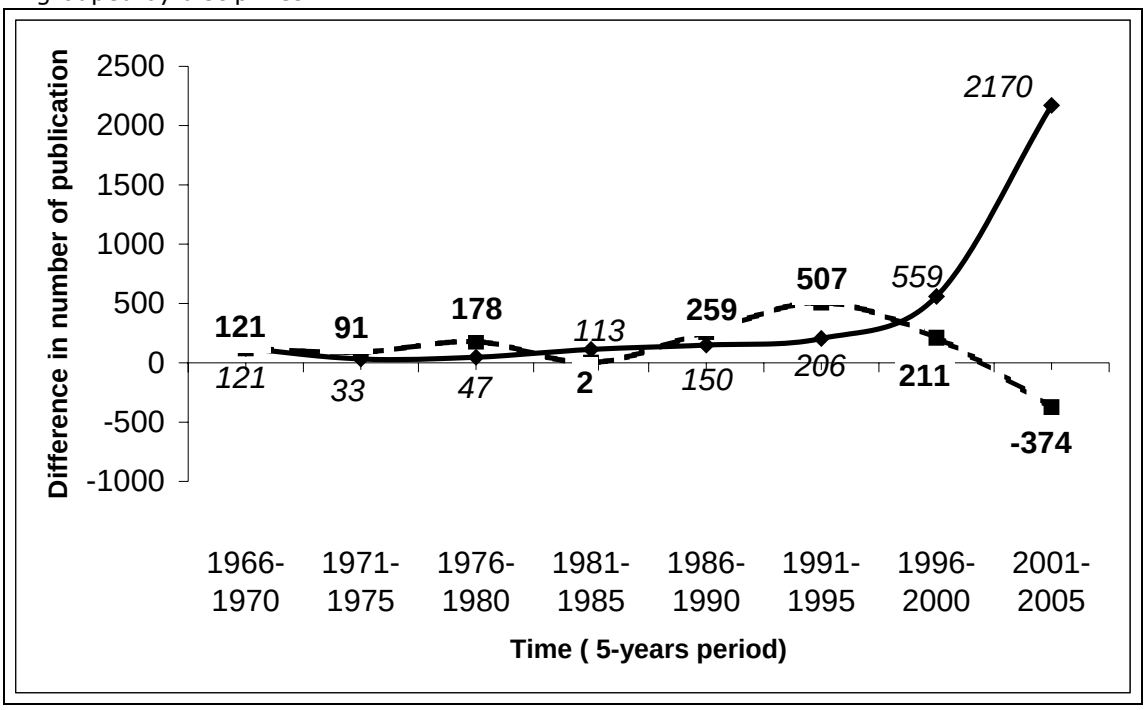


Figure 3. Differences in number of publication (Δ) from one period to the immediately previous one.

Figure 4. Differences in number of publication (Δ) from one period to the immediately previous one, grouped by disciplines**.**



Summary and implications

The present study takes advantage of bibliometric techniques to investigate the growth of literature on school-parents relationships, both in psychological databases and in non-psychological ones, evaluating also the use of different research strategies (queries) to gain access to the maximum number of relevant publications. To this end, about 11,000 records of publications ranging from year 1966 to 2005 have been taken into consideration.The results have a few direct and indirect implications on research in the field of parents-teachers relationships.

The analyses on the whole sample show a very strong increase in the growth during all the period. For each paper published in 1966, nowadays 35 papers are published.

For what regards distribution in time, it is possible to witness a concentration of works in the last 15 years of the period taken into consideration, when 77.8% of all the papers have been published. Narrowing the time span taken into consideration, 39.3% of the papers have been published in the last five years. Data thus show that studies in the relationship between parents and school are a rather recent phenomenon. This is coherent with the rapid spread, during the Nineties, of the principal models of parental involvement in schools.

The distribution of the principal search strings within the databases has shown that the use of strings like "parental involvement" and "family involvement" allows to retrieve 76.6% of all the papers. By the way, this also stress the relevance of the word "involvement" (and the consolidation of the research field around this concept). Moreover, this involvement is not limited to parents, but embraces many other possible actors in children's development.

For what regards the evaluation of growth curves in literature production, a linear model of growth can be excluded. The curve which best fits our data is probably a logistic curve, presently at its early stage of exponential growth.

For what regards differences between psychological and non-psychological literature, it is possible to draw some more conclusions (conclusions, of course, of a very provisional character, given the aforementioned problems regarding the segmentation of databases).

Many disciplines are involved in the field of relationships between parents and schools.
The number of publications in non-psychological fields is 5,256 vs 5,616 publications in psychological databases, an amount of research work of nearly the same size.

However, there are differences in the distribution in time of the contributions. In the last 5 years taken into consideration, psychological areas produced 58.4% of all the psychological work on this theme, but the non-psychological journals published only 18.9% of all the work in their area, showing that psychological journals tend to host more and more articles on the topic.

In non-psychological areas, in the last five years it is possible to record a decrease in publications with respect to the previous quinquennium (-374 publications). It is the first time from year 1966 that such a decrease happens. Together with the fact that in this same period the psychological journals published many more articles than in years 1995-2000 (+2.170), could this support the hypothesis that there is a trend towards a sort of "psychological specialization" of this field of studies?

References:

Baker, A.J.L. & Soden, L.M. (1998) *The challenges of parent involvement research*. ERIC Clearinghouse on Urban Education, ERCIC/CUE Digest n. 134, ED419030

De Solla Price, D. (1965) Networks of scientific paper. *Science*, 149(3683): 510-515.

Fan, X. & Chen, M. (2001) Parental involvement and students' academic achievement: a meta-analysis. *Educational Psychology Review*, 13 (1).

Fernandez-Cano, A., Torraldo, M. & Vallejo, M. (2004) Reconsidering Price's model of scientific growth: an overview. *Scientometrics*, 62(3):301-321.

Glanzel, W., Schoepflin, U. (1999) A bibliometric study of reference literature in the sciences and social sciences. *Information processing and management*, Vol.35: 31-44.

Gu, Y. (2004) Global knowledge management research: a bibliometric analysis. *Scientometrics*, 61(2)171-190.

Hicks, D.(1999) The difficulty of achieving full coverage of international social science literature and the bibliometric consequences. *Scientometrics*, 44(2):193-215.

Katz, J.S. (1999) *Bibliometric indicators and the social sciences*. Paper prepared for ESRC Polaris House. Retrieved November2008 from: http://www.sussex.ac.uk/Users/sylvank/pubs/ESRC.pdf

Katz, J. S., Hicks, D. (1997), Desktop Scientometrics. *Scientometrics*, 38 : 141--153.

Tague, J., Beheshti, J. & Rees-Potter, L. (1981) The laws of exponential growth: evidence, implications, and forecasts. *Library Trend*, 30(1) 125-149

Van Leeuwen, T. (2006) The application of bibliometric analyses in the evaluation of social science research. Who benefit from it, and why it is still feasible. *Scientometrics*, 66(1)133-154.

APPENDIX 1

Table 1.
Descriptions of databases

PsycINFO	"PsycINFO is an abstract database of psychological literature from the 1800s to the present. An essential tool for researchers, PsycINFO combines a wealth of content with precise indexing so you can get just what you need easily." (Retrieved nov. 2008 from: http://www.apa.org/psycinfo/)
Applied Social Sciences Index and Abstracts (ASSIA)	"Applied Social Sciences Index and Abstracts on the Web is an indexing and abstracting tool covering health, social services, psychology, sociology, economics, politics, race relations and education." (Retrieved nov. 2008 from: http://www.csa.com/factsheets/assia-set-c.php)
CSA Social Service Abstracts (CSA-SSA)	"CSA Social Services Abstracts provides bibliographic coverage of current research focused on social work, human services, and related areas, including social welfare, social policy, and community development." (Retrieved nov. 2008 from: www.csa.com/factsheets/supplements/ssaguide.pdf, p.3)
CSA Sociological Abstracts (CSA-SA)	"All subdisciplines of sociology are covered in CSA Sociological Abstracts. These include areas of study related to: Sociological theory, History of sociology, Urban sociology, Rural sociology, Sociology of marriage and family, Ethnic relations, Sociology of gender, Political sociology, Sociology of sports, Culture and social structure, Sociology of health and medicine, Sociology of crime. Content from such other disciplines as anthropology, social psychology, demography, education, criminology, penology, and political science is also included." (Retrieved nov. 2008 from: http://www.csa.com/factsheets/supplements/sakeyfacts.doc)
Educational Resources Information Center (ERIC).	"All materials included in ERIC must meet the selection standard of being directly related to the field of education. The collection scope spans early childhood education through higher education, vocational education, and special education; including teaching and teacher education, education administration, assessment and evaluation, counseling, information and technology, and the academic areas of reading, mathematics, science, and environmental education, languages, and social studies." (Retrieved nov. 2008 from: www.eric.ed.gov/ERICWebPortal/resources/html/news/ERIC%20Selection%20Policy.pdf)

APPENDIX 1

Table 3.
A synthesis of queries used for research and coding

Query	*Coding*	*Query*	*Coding*
Family involvement	fam_inv	Parents Involvement	prs_inv
Family participation	fam_par	Parents Participation	prs_par
Home-School connection	hs_con	Parent-Teacher Partnership	pt_par
Home-School partnership	hs_par	Parent-Teacher Relationship	pt_rel
Home-School relationship	hs_rel	Parent-Teacher Communication	pt_com
Parental involvement	par_inv	Home-School collaboration	hs_coll
Parental Participation	par_par		