A Meta-analysis of School Effectiveness Studies

Un metaanálisis de estudios de eficacia escolar

DOI: 10.4438/1988-592X-RE-2013-361-235

Jaap Scheerens
Bob Witziers
Rien Steen
Management. Enschede, The Netherlands.

Abstract

A meta-analysis was carried out on the basis of 155 research studies on school effectiveness, comprising a total of 1.211 associations between school effectiveness enhancing factors and student outcome variables. The original studies were carried out between 1984 and 2005. The school effectiveness enhancing conditions that were included were: cooperation between staff, an orderly school climate, monitoring, curriculum quality, homework, learning time, parental involvement, achievement orientation, educational leadership and differentiation. The outcome variables were student achievement results in Mathematics, mother tongue language and other subject matter domains, including Science. A multi level approach to meta-analysis was used, on the basis of which numerical effect sizes (Fischer's Z coefficients) were calculated. Effect sizes for the curriculum related factors, curriculum quality and teaching time were relatively the highest (,15), closely followed by the school climate factors (orderly climate and achievement orientation (,14). According to widely accepted standards (Cohen, 1969), these effects are to be considered as small effects; although in the discussion some arguments are provided that might lead to an upgrading of what one could call the "practical significance" of these effect sizes. The analyses included moderator variables representing study characteristics that were analyzed for their influence on the effect sizes of the various school effectiveness enhancing factors. For most factors, effect sizes were slightly higher for studies carried out in primary schools as compared to secondary schools. For other moderator variables, such as subject matter area, the results were less straightforward. Results are discussed for their

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

substantive educational implications, and their meaning for the field of empirical school effectiveness research.

Key words: school effectiveness, meta-analysis, educational effects, educational leadership, educational evaluation.

Resumen

Se llevó a cabo un metaanálisis de 155 estudios sobre eficacia escolar, que abarcó un total de 1.211 asociaciones entre factores de mejora de la eficacia escolar y variables de resultados de los alumnos. Los estudios originales se llevaron a cabo entre 1984 y 2005. Se incluyeron las siguientes condiciones para la mejora de la eficacia escolar: cooperación entre el personal, ambiente escolar ordenado, seguimiento, calidad curricular, deberes, tiempo de aprendizaje, implicación parental, orientación al logro, liderazgo educativo y diferenciación. Como variables de resultados se evaluaron los logros de los alumnos en Matemáticas, idioma materno y otras áreas temáticas, incluida la de Ciencias. Se realizó un enfoque multinivel de metaanálisis, sobre cuya base se calcularon los tamaños del efecto numéricos (coeficientes Z de Fischer). Los tamaños del efecto de los factores relacionados con el currículo, la calidad curricular y el tiempo de enseñanza eran relativamente los más elevados (0,15), seguidos de cerca por los factores relacionados con el ambiente escolar (ambiente ordenado y orientación al logro –0,14–). Según normas ampliamente aceptadas (Cohen, 1969), estos efectos deben considerarse pequeños; aunque en el debate se ofrecen algunos argumentos a favor de poner en valor lo que se podría llamar la significancia práctica de estos tamaños del efecto. En los análisis se incluyeron variables moderadoras relacionadas con características de estudio y se investigó su influencia en los tamaños del efecto de diversos factores de mejora de la eficacia escolar. La mayoría de los factores mostraron tamaños del efecto ligeramente superiores en los estudios realizados en escuelas primarias en comparación con escuelas secundarias. En otras variables moderadoras, como el área temática, los resultados fueron menos claros. Se discuten las implicaciones educativas sustantivas de los resultados, así como su significado para el campo de la investigación empírica sobre eficacia escolar.

Palabras clave: eficacia escolar, metaanálisis, efectos de educativos, liderazgo educativo, evaluación educativa.

Introduction. The field of educational effectiveness research

The elementary design of educational effectiveness research is the association of hypothetical effectiveness enhancing conditions of schooling and output measures, mostly student achievement. The major task of educational effectiveness research is to reveal the impact of relevant input characteristics on output and to "break open" the black box in order to show which process or throughput factors "work", next to the impact of contextual conditions. Among educational effectiveness studies a distinction can be made between studies that have concentrated on school level inputs and processes (school effectiveness studies) and studies on teaching at classroom level (instructional effectiveness studies). In more conceptual contributions authors have combined school and instructional effectiveness enhancing conditions in integrated educational effectiveness models (Bosker and Scheerens, 1995; Creemers, 1994; Stringfield and Slavin, 1995: Kyriakides, 2005, Creemers and Kyriakides, 2008).

Integrated educational effectiveness models have the following characteristics:

- Outputs are the basic criteria to judge educational effectiveness.
- In order to be able to properly evaluate output, achievement or attainment measures should be adjusted for prior achievement and other pupil intake characteristics; in this way the value added by schooling can be assessed.
- Multi-level structure, uniting effectiveness enhancing conditions at system, school, classroom and individual student level.

Research traditions in educational effectiveness vary according to the emphasis that is put on different kind of antecedent conditions of educational outputs. These traditions also have a disciplinary basis. The following research areas or research traditions have been considered in summarizing the research results (Scheerens and Bosker, 1997, ch. 2):

- Research on equality of opportunities in education and the significance of the school in this.
- Economic studies on education production functions.
- The evaluation of compensatory programs.
- Studies of unusually effective schools.

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

- Studies on the effectiveness of teachers, classes and instructional procedures.

An elaborate body of research studies and research reviews on these diverse strands of school effectiveness, over a period of three decades, is available from the literature. Early reviews are those by Anderson (1982), Cohen (1982), Dougherty (1981), Edmonds (1979), Good and Brophy (1986), Kyle (1985), Murnane (1981), Neufeld et ál. (1983), Purkey and Smith (1983), Ralph and Fenessey (1983), Rutter (1983), and Sweeney (1982). More recent reviews are those by Cotton (1995), Creemers (1994), Levine and Lezotte (1990), Reynolds et ál. (1993), Sammons et ál. (1995), Scheerens and Bosker (1997) and Teddlie and Reynolds (2000). Reviews of school effectiveness research in developing countries have been presented, among others, by Hanushek (1995), and by Fuller and Clarke (1994) –the latter review incorporates results of reviews by Fuller (1987), Lockheed and Hanushek (1988), Lockheed and Verspoor (1991)–. Other meta-analyses in this field are published by Marzano (2003), Creemers and Kyriakides, 2008 and Hattie, 2009.

More recent work would suggest the inclusion of perhaps two additional strands:

- Studies on failing schools.
- Comprehensive School Reform Programs.

Studies on failing schools relate work in school effectiveness research to school improvement approaches aimed at "turning around" failing schools (e.g.,. Stoll and Myers, 1997). More analytic and empirical contributions indicate that factors identified in school effectiveness research are also the school process dimensions on which failing schools are weak (Stringfield, 1994; Van der Grift and Houtveen, 2006).

Comprehensive School Reform Programs could be seen as implementations of integrated educational effectiveness models as they integrate aspects of school governance and management at school level and instructional approaches at classroom level. Their success has been demonstrated in reviews and meta-analysis, e.g., Borman, Carter, Aladjem and LeFloch (2004), Ross and Gil (2004), and Rowan, Camburn and Barnes (2004).

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

Results in terms of the factors that "work"

Reviews show considerable consensus in the range of factors that are seen as having received empirical support as malleable conditions of effective schooling. Based on their 1997 review, Scheerens and Bosker distinguish the following factors:

TABLE I. General effectiveness enhancing factors

1.	Achievement orientation / high expectations / teacher expectations
2.	Educational leadership
3.	Consensus and cohesion among staff
4.	Curriculum quality / opportunity to learn
5.	School climate
6.	Evaluative potential
7.	Parental involvement
8.	Classroom climate
9.	Effective learning time (classroom management)
10.	Structured instruction
11.	Independent learning
12.	Differentiation, adaptive instruction
13.	Feedback and reinforcement

Source: Scheerens and Bosker, 1997.

Scheerens and Bosker (1997) provide detailed definitions of each of these main factors, dividing them in sub-components. On the basis of a set of 72 more recent research articles, published in the period 1995-2005, this structure of main components and sub-components was further elaborated. The results are too extensive to be added as an appendix, but are available with the authors in an unpublished research report (Scheerens, Luyten, Steen and Luyten-de Thouars, 2007). The more elaborated analysis of questions, items and scales showed that the categorization as shown in Table 1, was useful as a framework to cover the more specific items and questions

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

Meta-analyses

Scheerens and Bosker (1997) presented the results of a meta-analysis on a subset of the factors that are listed in Table I. Results, expressed as the Fischer's Z coefficients about the association of the factor in question with an educational achievement measure, were as shown in Table II.

TABLE II. Results meta-analysis, studies before 1995

LABEL	Effect
School organizational factors	
Achievement pressure for basic subjects	0,14
Educational leadership	0,05
Monitoring / evaluation	0,15
Cooperation / consensus	0,03
Parental involvement	0,13
Orderly climate	0,11
Opportunity to learn	0,09
Time on task / homework	0,19/0,06
Monitoring at classroom level	0,11 (n.s.)

Source: Scheerens and Bosker, 1997, ch. 6.

According to the established conventions (Cohen, 1969) these coefficients should be considered as small effects[1].

In the remaining part of this article a meta-analysis is presented on 155 research studies, carried out between 1985 and 2005. These 155 research studies comprise of 1.211 replications, in the sense of associations of a variable representing a certain factor and an achievement outcome variable[2].

(1) Cohen refers to standardized effect sizes *d*, which are about twice the Fischer's Z coefficient; he indicates small effect sizes as about $d = ,20$ ($r = .10$) and medium effect sizes as about $d = ,50$ ($r = ,25$).

(2) This article is based on a larger study described in an internal report of the University of Twente.

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

The independent variables, in the sense of school effectiveness enhancing conditions, included in the meta-analyses, were 10 out of the 13 factors listed in Table II, namely: *consensus and cohesion among staff, orderly climate, monitoring, curriculum quality, homework, effective learning time, parental involvement, achievement orientation, educational leadership* and *differentiation.* With respect to the variable *school climate,* only replications covering *orderly climate* were used.

Methods

Literature search

A meta-analysis relies on collecting as many studies as possible regarding the topic of interest. The search methods included searches on the Web of Science, and the ERIC and ERA databases. The search was focused at articles published between 1995 and 2005. In addition, the literature database of ECER conferences was examined. In the search the following key words were used: school effectiveness, learning results, effectiveness, effective teaching, effective instruction, teacher effectiveness, educational effectiveness, school effectiveness, student achievement. Finally, recent reviews and books on school effectiveness were checked in order to find additional relevant literature ('snowball method').

The first step of this search resulted in several hundreds of publications. From these publications, about one-third appeared not to be useful for our purposes, while from one-sixth of all publications it could not be determined whether or not they contained useful information. These were articles that appeared to be inaccessible. This left us with 72 articles that contained information relevant for the purposes of our study. These articles were analyzed with regard to effect size presented on student achievement outcomes and relevant school effectiveness variables, while at the same time data were collected on particular study characteristics.

The resulting database was combined with an existing database concerning a meta-analysis on the same topic. This meta-analysis covered the period 1985-1995 and the results of this analysis were published by

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

Scheerens and Bosker (1997). For the goals of our research, this database was re-examined and, when necessary, information was added after consulting the original sources.

Meta-analysis

A multilevel approach to meta-analysis (Hox, 2002; Raudenbusch and Bryk, 1985) was applied. In this approach the selected studies are considered to be a sample from the population of studies, in our case this regards the relationship between specific school effectiveness indicators and student outcomes. Nested under each study are the secondary units: the schools. Each study can then be viewed as an independent replication. This concept could be used but would not solve the problem of biased estimates due to unidentified dependencies when applying multiple results from one study, e.g., when effects are reported for mathematics and language achievement in one study while using the same sample of schools and students. To deal with this problem, in stead of the two-level model for meta-analysis a three-level model was used, in which the highest level of the studies is referred to as the across-replication level, and the multiple results within a study as the within-replication level. The principal advantages of the statistical meta-analysis employed here are threefold: firstly, the information from each study is weighted by the reliability of the information, in this case the sample size and secondly, dependencies between within study replications are controlled for. Thirdly, the method applied enables us to examine which study characteristics (or moderators) are responsible for the variation in effect sizes.

To indicate the effect of school effectiveness variables, Fisher's Z transformation of the correlation coefficient was used. Not all studies presented their results in terms of correlations, and therefore all other effect size measures were transformed into correlations, using formulae presented by Rosenthal (1994). For small values of the correlation coefficient, Z_r and *r* do not differ much, but it should be remembered that all figures presented in the following and indicating effect sizes refer to Z_r.

Further details of the multi-level approach to meta-analysis that was used are given in the technical annex.

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

Results

Independent and dependent variables used in the meta-analysis

From the thirteen effectiveness enhancing conditions listed in Table II, 10 were included in our current meta-analysis. The selection was motivated by our intention to concentrate on variables that have a meaning at school level, despite of the fact that most of them also have an interpretation at class/teacher level. The variables from the list in Table I that we did not include are those that are intrinsically characteristics of instructional processes, namely *structured instruction*, *independent learning* and *feedback and reinforcement.*

Dependent variables used in our study were student outcomes in the cognitive domain, namely student achievement results in Mathematics, Language, and other subjects, including Science.

Moderator variables used in the meta-analysis

As it was stated in the above, our method allows us to model effect sizes as a function of study characteristics. A first relevant characteristic deals with the question of whether studies have used a language, a mathematics test score, or another score to assess student achievement. This moderator provides insight into the question as to which learning outcomes are most 'malleable' by school characteristics. Previous studies (Scheerens and Bosker, 1997) suggest that schools have more impact in the area of Mathematics than in the area of Language. In our study 45,3% of our data relate to the use of a math test, 33,8% of all results to a language test.

Apart from examining the impact of the type of test employed, we also investigated the effects of the country in which the study was conducted (the United States of America, the Netherlands, or other countries) and the education level or sector in which the study took place (Primary or Secondary Education). Results regarding these study characteristics provide insight into the question of which context is most 'susceptible' for school effectiveness indicators. Studies from the past show that, by and large, effect sizes are higher in US-schools and in primary schools (Scheerens and Bosker, 1997, chapter 6). In our study 33,5% of all effect sizes relate to studies conducted in the US, 24,5% to studies carried out in the Netherlands

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

and 42% to studies conducted in other countries. With regard to school type 63, 7% of all results relate to studies carried out in primary schools (36,3% in secondary schools).

The other moderator variables relate to the quality of the studies involved. One of them relates to the issue whether or not studies control for student intake characteristics. Effect sizes are by definition less accurate in case outcomes are not corrected for student intake characteristics. Almost all studies in our database include characteristics such as socio-economic status, age, gender, ethnicity and, in a minority of cases, prior achievement, implying that only in rare cases the dependent variable represents learning gain. We therefore included "multi-level/not multi-level" as an additional moderator variable in our analyses. In our study 57,6% of all results are based upon multi-level techniques, the other 42,4% on other techniques.

Finally, most of our independent variables have a meaning at school as well as at classroom level and for these variables an additional moderator was included in the analysis. The moderator in question represents the level at which the school effectiveness indicator of interest was measured; was the indicator measured at the class or school level? An example concerns the analysis of data relating to the concept of *monitoring.* This indicator is sometimes measured at the school level (for example by investigating whether a monitoring system is used by the school) and sometimes at the class level (for example by checking the amount of time spent by a teacher in monitoring pupil's progress). The second example is school climate versus classroom climate. For the indicators *consensus and cohesion among staff, parental involvement, educational leadership* and *differentiation* only school level information was used, so that the "level" moderator was not applied to these.

Results: multi-level approach

The results of the multi-level approach to meta-analysis are presented in Table III and Table IV. Table III shows the average effect sizes for all independent variables, and results are generalized over all moderator.

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

TABLE III. Multi-level, empty model

	NUMBER OF CASES		VARIANCE		
	Across replications	Within replications	Mean effect size	Across replications	Within replications
Consensus and cohesion among staff (cooperation)	28	83	0,019	0,001	0,000
Orderly climate	46	170	0,129 ***	0,026 *	0,008 **
Monitoring	43	194	0,061 ***	0,003	0,021 **
Curriculum quality; OTL	25	43	0,145 ***	0,028	0,007
Homework	21	56	0,073 **	0,019	0,000
Effective learning time	30	111	0,147 ***	0,014 **	0,017
Parental involvement	42	142	0,093 ***	0,018 ***	0,000
Achievement orientation	50	135	0,141 ***	0,036 ***	0,010
Educational leadership	53	170	0,046 *	0,025 *	0,000
Differentiation	30	107	0,017	0,021 ***	0,008

TABLE IV. Multi-level, with moderators as predictors

	Intercept	Secondary	Arithmetic/ Math	Language	USA	The Netherlands	Value added	Not multilevel	Class/teacher level	VARIANCE	
										Across replications	Within replications
Consensus and cohesion among staff (cooperation)	-0,058 ***	0,065 ***	-0,006	0,004	0,053 ***	0,031 *	0,032 **	0,004	-	0,000	0,000
Orderly climate	0,135	0,025	-0,004	0,001	0,082	0,005	-0,076	0,017	-0,009	0,023 **	0,008 **
Monitoring	0,121 **	-0,070 **	0,106 *	0,117 **	-0,116 **	-0,098 ***	-0,084	0,088 ***	-0,052 **	0,001	0,019 **
Curriculum quality; OTL	-0,047	-0,066	0,151	0,008	-0,047	-0,007	0,110	0,199	0,013	0,012 **	0,006 **
Homework	0,233	-0,055	0,036	0,020	0,335 ***	0,376 ***	-0,410 ***	0,087	-0,166 **	0,000	0,000
Effective learning time	0,191 ***	-0,185 **	-0,039	0,058	-0,092	-0,145 *	0,039	0,210 ***	-0,090	0,002	0,019 *
Parental involvement	0,213 ***	-0,005 ***	-0,010	-0,013	0,114 *	-0,028	-0,136 **	-0,105	-	0,019 ***	0,000
Achievement orientation	0,202 ***	-0,063	-0,022	-0,027	0,070	-0,154 **	-0,180	0,000	0,047	0,028 ***	0,010
Educational leadership	0,052	-0,002	0,009	0,011	0,050	-0,095	0,012	-0,018	-	0,022 **	0,000
Differentiation	-0,085	-0,067	-0,035	-0,007	0,245 **	0,052	0,046	0,232 ***	-	0,006 **	0,007

Effect-sizes marked as (*) are significant at the 0,10 level; those marked as (**) are significant at the 0,05 level; and those marked as (***) are significant at the 0,001 level.

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

Cooperation

The results of the meta-analysis for the factor *consensus and cohesion among staff (cooperation)* and its impact on pupil achievement show that in total 28 studies were included, some of which contained multiple results, leading up to a total number of 83 within replications.

The estimated mean effect size of cooperation across all studies equals a Fischer's Z value of 0,019. The estimated variance across all studies (both within and across replications) is ,001. This indicates that the 95% prediction interval around the means ranges between Z_r = -0,043 and Z_r = 0,081.

The prediction interval, in contrast to the confidence interval, describes the distribution based on the estimates. The confidence interval gives only information on the degree of precision with which the mean of that distribution is estimated.

The results of the analysis trying to predict differences between effect sizes with moderators such as subject matter, sector, study design and others indicate that some of the moderators have a significant relationship with the effect size. Studies carried out in the USA and The Netherlands, studies carried out in secondary schools and studies employing a value added design show significantly higher effect sizes. These results do not change the overall conclusion that cooperation among teachers appears to be an insignificant variable in explaining variation in pupil achievement. For example, controlling for other study characteristics, US-studies have an average effect size around zero (-0,058 + 0,053). For the Netherlands this figure is -0,027 (-0,058 + 0,031).

Orderly climate

The estimated mean effect size of orderly climate is ,129, which is significant at the 1% level. The mean effect size is based on in total 46 studies, most of them with multiple results. The total of within replications is 170. The estimated variation across all studies (both across and within replications) equals 0,034. This indicates that the 95% prediction interval around the mean effect size is between Z_r = -0,231 and Z_r = 0,489.

The results of the analyses trying to predict differences between effect sizes show that none of the moderators has a significant relationship with the mean effect size. This means, for example, that there is no difference between studies measuring this concept at the school level or studies measuring it at the class level; the effect size in both types of studies is equal.

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

Monitoring

The estimated effect size of monitoring across all 25 studies involved in our analyses is $Z_r = 0.06$ which is significant at the 1% level. The estimated variance across all studies (both within and across replications) is 0,024, indicating that the 95% prediction interval around the mean effect size runs from $Z_r = -0{,}243$ to $Z_r = 0{,}3625$ The results of the analyses trying to establish relationship between effect sizes with study characteristics show that many moderators have a significant relationship with the effect size. Some of them have comparably a rather strong positive relationship with the effect size, others a rather strong negative relationship. For example, when a language test is the outcome variable, the effect size is about 0,12 higher, and when an arithmetic or mathematics test is used the effect size is about 0,11 higher, than in cases where other outcome variables were used. This implies that the mean effect size for monitoring is strongly diminished when outcomes in other subject are included, whereas effect sizes for Language and Mathematics are relatively high, 0,18 and 0,17, respectively. A smaller difference (0,09) in effect size is noted with respect to studies that do not use multi level modelling, as compared to studies that do use multi-level modelling. On the other hand, effect sizes turn out lower for studies carried out in the USA compared to all other countries. The same results apply to studies carried out in The Netherlands. On average, effect sizes for studies carried out in The Netherlands and the USA are around zero. When monitoring is measured at school level it has a higher effect size than in cases where it is measured at classroom level (0,05 higher). Finally, studies conducted in secondary schools show lower effect sizes than studies carried out in primary education (difference 0,07).

Curriculum quality

The curriculum quality concept includes three variables; opportunity to learn, effective learning time and homework.

The analysis concerning *opportunity to learn* involves 25 studies with 43 results in total. The mean effect size is 0,145 which is significant at the 1% level. The 95% prediction interval around the means ranges between $Z_r = -0{,}222$ and $Z_r = 0{,}512$. The analysis reveals further that there is hardly any variance among studies with regard to their effect sizes. Not surprisingly, there are no significant relationships between the moderators and the effect size.

With regard to *homework* 21 studies were analyzed involving a total of 56 results. The estimated mean effect size is 0,073. This indicates that the

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

95% prediction interval ranges between Z_r = -0,197 and Z_r = 0,343. Although the variance in effect appears relatively small, there are important differences between studies and the effect sizes they yield. This regards, first of all differences between countries. Studies conducted in both the US and The Netherlands yield effect sizes which are much higher than the effect sizes yielded by studies carried out in other countries. US-studies have an effect size which differs 0,345 from all other studies, while Dutch studies differ 0,376 from all other studies. Moreover, studies employing multi-level techniques produce much lower effect sizes. Finally, studies measuring the concept at the class level (in fact 98% of all studies) yield significantly lower effect size than studies measuring the concept of homework at the school level.

The estimated effect size of *effective learning time* equals 0,147 (significant at the 1% level). This indicates that the 95% prediction interval ranges between Z_r = -0,0197 and Z_r = 0,491. The analysis relating moderators to the effect size indicate that studies carried out in primary schools show significantly lower effect sizes (0,19), while studies employing other than multi-level techniques yield significantly higher effect sizes (a difference of 0,21). Finally, there is also a difference between countries. Studies carried out in The Netherlands come up with significantly lower effect sizes (-0,145). On average, the effect size of Dutch studies is about 0,05 (0,191-0,145).

Parental involvement

The analyses concerning parental involvement involve 42 studies, again with most of them having multiple results. In total there are 142 replications within the studies. The estimated effect size of parental involvement in all studies is Z_r = 0,093, which is significant at the 1% level. The 95% prediction interval around the means ranges between Z_r = -0,169 and Z_r = 0,355.

The data also show significant variation in effect sizes. The most important moderators in this respect are, respectively, whether or not the study involved controls for student characteristics affecting learning achievement and the country in which the study has been carried out. Not surprisingly, with regard to the former, studies taking into account student characteristics show significantly lower effect sizes (difference in coefficient of -0,14). With regard to the latter, effect size of studies carried out in the US are significantly higher than effect sizes of studies carried out in all other

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

countries (difference in coefficient of 0,11); controlling for the impact of other moderators, the effect size in US-studies is on average 0,327 (0,213 + 0,114).

Achievement orientation

The estimated mean effect size for achievement orientation is 0,147, which is significant at the 1% level. The 95% prediction interval around the means ranges between Z_r = -0,279 and Z_r = 0,561. The figures presented are based on 30 studies containing 81 results.

Once again the data indicate that there is significant variation in effect sizes across studies. However, only one moderator is of significance in this respect. Studies conducted in The Netherlands have significantly lower effect sizes than studies carried out in other countries (-0,15). Moreover, an interesting fact is that it does not seem to matter at which level this concept is measured. There is no significant difference between studies measuring this concept at the school level and studies measuring this concept at the class level.

Educational leadership

Another frequently studied school effectiveness indicator is educational leadership (53 studies with 170 results). The mean effect size in this case is 0,046. This figure is significant at the 10% level. The 95% prediction interval around the means ranges between Z_r = -0,263 and Z_r = 0,355. The analyses cannot detect any significant relationship between the moderators distinguished in this study and the effect size, although studies' effect sizes vary significantly around the mean.

Differentiation

The last concept investigated in this study is differentiation. The analysis concerning this concept involves 30 studies with in total 107 different results. The mean effect size found is 0,017, a figure which does not deviate significantly from zero. The 95% prediction interval lies between Z_r = -0,317 and Z_r = 0,351. The results of the analysis examining the variation in the effect size show that two moderators are important. Effect sizes are significantly higher in US-studies than in studies conducted in all other countries (a difference of 0,245) and in studies using other techniques than multi-level techniques (a difference of 0,232).

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

Conclusion

Over viewing our quantitative results the conclusion is that, in general, the effect sizes found in our analysis range between 0,017 and 0,147. In terms of Cohen's *d* (which is approximately twice the size of the correlation coefficient) this means that the results vary from negligible to small.

In this respect they resemble the results of a previous meta-analysis presented by Scheerens and Bosker (1997). The results are also similar in the sense that the effect sizes found in this study for the different effectiveness indicators are comparable to the ones found previously. The biggest differences are found with respect to *parental involvement* (now ,09; then ,13), *effective learning time* (now about 015; then 0,19), *monitoring* (now 0,06; then 0,14) and *curriculum quality (opportunity to learn)* (now 0,13; then 0,08). The conclusion with respect to monitoring should be modified, however, since effect sizes in important subject matter areas as language and mathematics are in the order of 0,18 and 0,17 in our current analysis.

With respect to the impact of the moderator variables, our results indicate that, as was expected, for practically all variables, effect sizes are smaller when outcomes are adjusted for student background characteristics, and for all but two variables effect sizes are smaller when multi-level analyses are applied. There is also a relatively consistent slightly higher effect size for studies carried out in primary, as compared to studies conducted in secondary schools. The picture is less clear-cut for the moderator variables subject matter area and country (Table II).

When comparing our results to those found in the meta-analyses by Hattie (2009) and by Creemers and Kyriakides (2008), we see that these authors found effect sizes for educational readership of ,18 and ,07, respectively, for monitoring and evaluation ,31 and ,18 respectively and for orderly climate, ,17 and ,12. The effect sizes reported by Creemers and Kyriakides are quite similar to ours, while those reported by Hattie are higher. According to Hattie (2009, p. 202) this might be caused by the fact that stricter quality controls were used in selecting studies in the Europe-based meta-analyses.

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

Discussion

According to Cohen's standards for interpreting effect sizes, our results on school effectiveness indicators should be interpreted as negligible to small. It should be noted however, that several authors argue that Cohen's standards are to be considered as too conservative, and do not match the practical significance of malleable school variables. Richard, Bond and Stokes-Zoota (2003; cited by Baumert et ál., 2006) found a mean correlation of r = ,21 in their meta-analysis of meta-analyses in social psychology, and proposed a modification of Cohen's classification, considering a correlation of ,30 to indicate a large effect (p. 339). Baumert, Luedtke and Trautwein (2006) propose the learning gain during one school year as a realistic standard to express effects of schooling. They cite several studies that indicate that this learning gain has the magnitude of about d = ,30. These authors also discuss a method to compute effect sizes developed by Tymms, Merrell and Henderson (1997), which, when applied to a practical example, suggests that effect sizes of about r = ,15 to ,20 (small to medium, according to Cohen's standards) would equal the learning gain in one school year, which they consider an effect of huge practical relevance. Seen in this light the effect sizes that we found for a number of school effectiveness indicators (in particular *school climate, curriculum quality, learning time* and *achievement orientation*) should be upgraded in their rating for practical significance.

Among the set of school effectiveness indicators that were studied the curriculum related and climate related factors showed the largest effects. *Opportunity to learn* and *learning time* had effect sizes of 0,15; whereas *orderly climate* and *achievement orientation* had effect sizes of 0,13 and 0,14, respectively. The relative importance of the curriculum variables underlines the importance of the content dimension in schooling. The time factor is interpreted in the sense of the temporary engagement with content, and, in this way, as a dimension of the implemented curriculum. The results on *homework* can be given a similar interpretation, where the effect size for homework was ,07. The realization that content and exposure to content matters could be interpreted as supporting the view that pro-active structuring of content, as in externally developed curricula and lesson plans, has a rightful place among school improvement strategies. This result speaks to the debate concerning school based, "bottom up" school improvement strategies versus the implementation of

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

external curricula. The former approach has been the preferred approach among scholars in the field of educational change (cf. Miles, 1998) but has been criticized, among others by Slavin (2000), and Muijs and Reynolds (2001) who describe the bottom up approach as "the ownership paradigm", in which the "re-invention of the wheel" by individual schools is put down as an inefficient approach. A similar line of argumentation, favoring externally developed curriculum material, is used with respect to the approach followed in Comprehensive School Reform Programs in the USA (Borman et ál., 2004).

The relatively high effect sizes concerning an *orderly school climate* are in line with results from large scale international assessment studies, like OECD's PISA program. More in depth analyses of these results (Luyten, Scheerens, Visscher, Maslowski, Witziers and Steen, 2005), however, indicated that the climate effects were heavily confounded with school composition, in the sense of school average socioeconomic status *(ses)* of the students. More specifically these results showed that schools with a better climate were more likely to have higher level *ses* composition.

The second climate factor, *achievement orientation* is based on variables like: clear focus on mastering basic subject, high expectations of students' achievement, and record keeping of students' achievement. High expectations reflect an active, optimistic attitude that seeks to get the best out of all students, and is related to the personality characteristic of internal locus of control. At the same time measures of high expectations might express a more reactive attitude, in which relatively high achievement is more like a cause, rather than an effect of high expectations.

Two variables that should be considered of high policy relevance in effective schooling, *monitoring* and *educational leadership* came out as having very small average effect sizes (0,06 and 0,05 respectively). The evaluation and feedback mechanism is considered as a promising lever for organizational learning and school improvement, an expectation that is at least reasonably met for Language and Mathematics outcomes (effect sizes of 0,18 and 0,17, respectively). Scheerens and Bosker (1997) report an average effect size for monitoring of 0,15. In evaluation studies concerning types of school evaluation and monitoring results show a mixed pattern as well. Schildkamp (2007) reports relatively disappointing results of evaluations of school self-evaluation programs (Schildkamp, 2007). Research results on the impact of system level accountability policies (Carnoy et ál., 2003), however, indicate that the combination of a high

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

internal evaluative potential of schools and a context of high stakes external accountability policy is effective in enhancing student performance.

The effect size for *educational leadership* (0,05) confirms a similar effect size as reported in Scheerens and Bosker (1997). More in depth analysis by Witziers, Bosker and Krüger (2003) focussed on indirect effects of educational leadership, where the interesting question is the one about the identification of variables that mediate the effect of leadership. Their results, and those of later studies, provide little consistency between studies, concerning the intermediary variables that were identified. The table below provides an overview.

TABLE V. Intermediary variables in studying indirect effects of school leadership

Reference of study	Significant intermediary variables
Hallinger and Heck, 1998	Learning climate Principal's instructional efforts
Hallinger, Bickman and Davis, 1996	A clear school mission Students' opportunity to learn Teachers' expectations
Hill, Rowe and Holmes-Smith, 1995	Teacher student interactions Professional climate
Bosker, De Vos and Witziers, 2000	Teachers' job satisfaction Teachers' achievement orientation Evaluation and feedback practices
Kythreotis and Pashiardis, 2006	Teachers' commitment to the school Teachers' academic emphasis Personal achievement goal orientations Classroom performance-goal structure

The school effectiveness indicator *cooperation and consensus* is a factor that makes perfect practical sense and has an important place in conceptual models of school effectiveness (e.g. the model developed by Creemers; cf. Creemers and Kyriakides, 2006). It comes out weak in our current and previous meta-analyses ,02 and ,03 respectively. This low effect may be due to the rather superfluous way in which this variable is often measured, for example in terms of the frequency of staff meetings.

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

Parental involvement had a small effect (0,09) as compared to Cohen's categorization, but, given the consideration mentioned at the beginning of this discussion, this might still be of practical significance. *Differentiation* had a negligible effect (0,02), but it should be noted that this factor, measured at school level refers to school level policies, and would potentially have higher impact when studied at classroom level.

A final issue, of a more methodological nature, to be raised has to do with strengths and weaknesses of a meta-analysis in which the effect of each relevant factor is estimated separately. The main advantage is that this approach attempts to show what each and every factor is 'worth' in its association with student achievement. A major disadvantage is the fact that the approach does not take the inter-correlations between the factors, nor relevant contextual variables, nor intermediary variables (particularly classroom level instructional processes) into consideration. In this sense, in order to give a more complete overview of the knowledge base from school effectiveness research, additional review of studies that have examined more complex configurations of the factors that were dealt with as discrete, independent factors in our analysis, is needed. Four kinds of studies should be mentioned, that would be complimentary to our approach:

- Studies that investigate indirect effects; the examples that were presented in the above, concerning educational leadership, illustrate this approach.
- Studies that have attempted to model alternative specifications of across level (e.g. school, classroom, individual student) relationships (e.g. Bosker and Scheerens, 1995; Hofman, 1995; Reezigt, Guldemond and Creemers, 1999). Such studies are sparse and their potential usually limited by data constraints.
- Studies that are driven by elaborate conceptual models of school effectiveness, e.g. Creemers and Kyriakides, 2006; Kyriakides, 2005. The same qualification applies as for the previous category.
- Evaluations and meta-evaluations of Comprehensive School Reform Programs, which include all factors that were studied in our meta-analyses (and more; comprising also instructional variables and above school policies). The positive outcomes of these studies, (e.g. Borman et ál., 2004) present probably the most robust empirical support for the "effective school model", so far. The more so,

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

because the quasi experimental design of the evaluations is one step further in allowing for causal interpretation of the research findings, than is the case for the mostly survey based, non experimental nature of the typical research study on which our analyses are based.

Bibliographic References

Anderson, C. S. (1982). The Search for School Climate: A Review of the Research. *Review of Educational Research,* 52 (3), 368-420.

Baumert, J., Lüdtke, O. y Trautwein, U. (2006). *Interpreting Effect Sizes in Large-Scale Educational Assessments.* Berlin: Max Planck Institute for Human Development.

Borman, K. M., Carter, K., Aladjem, D. K. y LeFloch, K. C. (2004). Challenges for the Future of Comprehensive School Reform. In C. T. Cross (Ed.), *Putting the Pieces Together: Lessons from Comprehensive School Reform Research,* (150). Washington D. C.: The National Clearinghouse for Comprehensive School Reform.

Bosker, R. J. y Scheerens, J. (1995). A Self-Evaluation Procedure for Schools Using Multilevel Modelling. *Tijdschrift voor Onderwijsresearch,* 20 (2), 154-164.

Bosker, R. J., Vos, H. de y Witziers, B. (2000). *Theories and Models of Educational Effectiveness.* Enschede: University of Twente, Department of Educational Organization and Management.

Bosker, R. J. y Witziers, B. (1995). *A Meta Analytical Synthesis of School Effectiveness Research: The Size of School Effects and the Effect Size of Educational Leadership.* Enschede: University of Twente, Department of Educational Organization and Management.

Carnoy, M., Elmore, R. y Siskin, L. (Eds.). (2003). *The New Accountability. High Schools and High-Stakes Testing.* New York; London: Routledge Falmer.

Cohen, M. (1982). Effective Schools: Accumulating Research Findings. *American Education,* January-February, 13-16.

Cohen, J. (1969). *Statistical Power Analysis for the Behavioural Sciencescational* (2nd ed.). Hillsdale (New Jersey): Lawrence Erlbaum and Associates.

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

Cotton, K. (1995). *Effective Schooling Practices: A Research Synthesis. Update. School Improvement Research Series.* Northwest Regional Educational Laboratory.

Creemers, B. P. M. (1994). *The Effective Classroom.* London: Cassell.

Creemers, B. y Kyriakides, L. (2006). Critical Analysis of the Current Approaches to Modelling Educational Effectiveness : The Importance of Establishing a Dynamic Model. *School Effectiveness and School Improvement,* 17 (3), 347-366.

— (2008). *The Dynamics of Educational Effectiveness: a Contribution to Policy, Practice and Theory in Contemporary Schools.* London: Routledge.

Dougherty, K. (1981). After the Fall: Research on School Effects since the Coleman Report. *Harvard Educational Review,* 51, 301-308.

Edmonds, R. R. (1979). Some Schools Work and More and More Can. *Social Policy,* 9, 28-32.

Fuller, B. (1987). What Factors RAISE achievement in the Third World? *Review of Educational Research,* 57, 255-292.

— y Clarke, P. (1994). Raising School Effects while Ignoring Culture? Local Conditions and the Influence of Classroom Tools, Rules and Pedagogy. *Review of Educational Research,* 64, 119-157.

Good, Th. L. y Brophy, J. (1986). School Effects. En M. C. Wittrock (Ed.), *Handbook of Research on Teaching,* (328-375). New York: McMillan.

Grift, W. J. C. M. van de y Houtveen, A. A. M. (2006). Underperformance in Primary Schools. *School Effectiveness and School Improvement,* 17 (3), 255-274.

Hallinger, P., Bickman, L. y Davis, K. (1996). School Context, Principal Leadership, and Student Reading Achievement. *The Elementary School Journal,* 96 (5), 527-550.

Hallinger, Ph. y Heck, R. H. (1998). Exploring the Principal's Contribution to School Effectiveness: 1980-1995. *School Effectiveness and School Improvement,* 9 (2), 157-191.

Hanushek, E. A. (1995). Interpreting Recent Research on Schooling in Developing Countries. *The World Bank Research Observer,* 10 (227-246).

Hattie, J. A. C. (2009). *Visible Learning: A Synthesis of over 800 Meta-Analyses relating to Achievement.* London: Routledge.

Hofman, A. W. H. (1995). Cross-Level Relationships within Effective Schools. *School Effectiveness and School Improvement,* 6, 146-174.

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

Hox, J. (2002). Multilevel Analysis: Techniques and Applications. Mahwah: Lawrence Erlbaum Associates.

Kyle, M. J. (Ed.). (1985). *Reaching for Excellence. An Effective Schools Sourcebook.* Washington D. C.: US Government Printing Office.

Kyriakides, L. (2005). Extending the Comprehensive Model of Educational Effectiveness by an Empirical Investigation. *School Effectiveness and School Improvement,* 16 (2), 103-152.

Kythreotis, A. y Pashiardis, P. (2006). The Influences of School Leadership Styles and Culture on Students' Achievement in Cyprus Primary Schools. San Francisco: AERA.

Levine, D. K. y Lezotte, L. W. (1990). *Unusually Effective Schools: A Review and Analysis of Research and Practice.* Madison: Nat. Centre for Effective Schools Research and Development.

Lockheed, M. y Hanushek, E. (1988). Improving Educational Efficiency in Developing Countries: What Do We know? *Compare,* 18 (1), 21-38.

Lockheed, M. y Verspoor, A. (1991). *Improving Primary Education in Developing Countries.* London: Oxford University Press.

Luyten, J. W., Scheerens, J., Visscher, A. J., Maslowski, R., Witziers, B. y Steen, R. (2005). *School Factors related to Quality and Equity. Results from PISA 2000.* Paris: OECD.

Marzano, R. J., (2003) *What Works in Schools. Translating Research into Action.* Alexandria: Association for Supervision and Curriculum Development.

Murnane, R. J. (1981). Interpreting the Evidence on School Effectiveness. *Teachers College Record,* 83, 19-35.

Neufeld, E., Farrar, E. y Miles, M. B. (1983). A Review of Effective Schools Research: The Message for Secondary Schools. Washington D. C.: National Commission on Excellence in Education.

Purkey, S. C. and Smith, M. S. (1983). Effective Schools: a Review. *The Elementary School Journal,* 83 (4), 427-452.

Ralph, J. H. y Fennessey, J. (1983). Science or Reform: some Questions about the Effective Schools Model. *Phi Delta Kappan,* 64 (10), 689-695.

Raudenbush, S. W. (1994). Random Effects Models. En H. Cooper y L. V. Hedges (Eds.), *The Handbook of Research Synthesis,* (301-321). New York: Russell Sage Foundation.

Raudenbush, S. W. y Bryk, A. S. (1986). A Hierarchical Model for Studying School Effects. *Sociology of Education,* 59, 1-17.

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

Reezigt, G. J., Guldemond, H. y Creemers, B. P. M. (1999). Empirical Validity for a Comprehensive Model on Educational Effectiveness. *School Effectiveness and School Improvement,* 10 (2), 193-216.

Reynolds, D., Hopkins, D. y Stoll, L. (1993). Linking School Effectiveness Knowledge and School Improvement Practice: towards a Synergy. *School Effectiveness and School Improvement,* 4 (1), 37-58.

Richard, F. D., Bond, C. F., Jr., y Stokes-Zoota, J. J. (2003). One Hundred Years of Social Psychology Quantitatively Described. *Review of General Psychology,* 7, 331-363.

Rosenthal, R. (1994). Parametric Measures of Effect Size. En H. Cooper y L. V. Hedges (Eds.), *The Handbook of Research Synthesis,* (231-244). New York: Russell Sage Foundation.

Ross, S. M. y Gil, L. (2004). The Past and Future of Comprehensive School Reform: Perspectives from a Researcher and Practitioner. En C. T. Cross (Ed.), *Putting the Pieces Together: Lessons from Comprehensive School Reform Research,* (151-174). Washington D. C.: The National Clearinghouse for Comprehensive School Reform.

Rowan, B., Camburn, E. y Barnes, C. (2004). Benefiting for Comprehensive School Reform: A Review of Research on CSR Implementation. En C. T. Cross (Ed.), *Putting the Pieces Together: Lessons from Comprehensive School Reform Research,* (1-52). Washington D. C.: The National Clearinghouse for Comprehensive School Reform.

Rutter, M. (1983). School Effects on Pupil Progress: Research Findings and Policy Implications. *Child Development,* 54 (1), 1-29.

Sammons, P., Hillman, J. y Mortimore, P. (1995). *Key Characteristics of Effective Schools: A Review of School Effectiveness Research.* London: OFSTED.

Scheerens, J. y Bosker, R. J. (1997). *The Foundations of Educational Effectiveness.* Oxford: Elsevier Science Ltd.

Scheerens, J., Luyten, H., Steen, R. Y. y Luyten-de Thouars (2007). *Review and Meta-Analyses of School and Teaching Effectiveness.* Enschede: University of Twente, Department of Educational Organisation and Management.

Schildkamp, K. (2007). *The Utilization of a Self-Evaluation Instrument for Primary Education.* (Dissertation). University of Twente, Department of Educational Organization and Management, Enschede, The Netherlands.

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

Stoll, L. y Myers, K. (1997). *No Quick Fixes. Perspectives on Schools in Difficulty.* London: Falmer Press.

Stringfield, S. (1994). Outlier Studies of School Effectiveness. En D. Reynolds et ál. (Eds.), *Advances in School Effectiveness Research and Practice.* Oxford: Pergamon.

Stringfield, S. C. y Slavin, R. E. (1992). A Hierarchical Longitudinal Model for Elementary School Effects. En B. P. M. Creemers y G. J. Reezigt (Eds.), *Evaluation of Educational Effectiveness,* (35-39). Groningen: ICO.

Sweeney, J. (1982). Research Synthesis on Effective School Leadership. *Educational Leadership*, 39, 346-352.

Teddlie, C. y Reynolds, D. (2000). *The International Handbook of School Effectiveness Research.* London: Falmer Press.

Tymms, P., Merrell, C. y Henderson, B. (1997). The First Year at School. A Quantitative Investigation of the Attainment and Progress of Pupils. *Educational Research and Evaluation,* 3 (2), 101-118.

Witziers, B., Bosker, R. J. y Krüger, M. (2003). Educational Leadership and Student Achievement: the Elusive Search for an Association. *Educational Administration Quarterly,* 39 (3), 398-425.

Dirección de contacto: Jaap Scheerens. University of Twente. Faculty of Behavioural Sciences. Department of Educational Organization and Management. PO Box 217; 7500 AE Enschede, The Netherlands. E-mail: j.scheerens@utwente.nl

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

Technical annex. Statistical modeling issues

The multilevel model for the meta-analysis is:

$$\delta_{rs} = \delta_0 + \gamma_1 \text{ subject-math}_{rs} + \gamma_2 \text{ subject-lang}_{rs} + \gamma_3 \text{ sector}_s + \gamma_4 \text{ country-USA}_s + \gamma_5 \text{ country-NL} + \gamma_6 \text{ design}_s + \gamma_7 \text{ statistical technique employed}_s + \gamma_8 \text{ level}_s + u_{rs} + v_s + e_{rs} ( )$$

where:
δ_{rs} is the effect size d in replication r in study s, which is an estimate of the population parameter δ_{rs}
e_{rs} is the associated sampling error with δ_{rs}
δ_0 is the effect size across studies
v_s is the associated sampling error with δ_0 (the across replications sampling error)
u_{rs} is the associated sampling error with δ_s (the within replications sampling error)
g_1 through g_8 are coefficients with the following predictors:

subject-math	0 = not math only, 1 = math only
subject-lang	0 = not language only, 1 = language only.
sector	0 = primary education, 1 = secondary education
country-USA	0 = not USA, 1 = USA
country-NL	0 = not The Netherlands, 1 = The Netherlands
design	0 = gross, 1 = value added (correction for prior achievement and/or background variables)
statistical technique employed	0 = multilevel technique, 1 = not multilevel technique
level	0 = teacher/class level, 1 = school/school leader level

Thus in the equation δ_0 is the estimated effect size for studies where all predictors have value 0.

The model is the same as used by Bosker and Witziers (1995), which is based on a model from Raudenbush (cf. Raudenbush and Bryk, 1986; Raudenbush, 1994).

Note:
Moderator *subject* consists of three categories, i.e. 'math only', 'language only' and a rest category, mainly containing composite scores and subject Science. Because only binary variables can be handled in the above equation, for subject-math, category 0 contains both 'language' and the rest category. Likewise for subject-lang, category 0 contains both 'math' and the rest category.

Also moderator *country* consists of three categories, i.e. 'USA', 'The Netherlands' and 'countries not being USA or The Netherlands'. Where in the text a comparison is made between USA and 'other countries', these other countries include the Netherlands. Likewise, where the Netherlands are compared with 'other countries', these other countries include the USA.

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012

Effect sizes

In the studies from which the results are analysed here, the effect sizes have been reported in various ways. Most of the effect sizes could be transformed directly into Fisher's Z using formulae presented by Rosenthal (1994).

Where in one original analysis (e.g. a ML analysis) two or more indicators of the same concept were used, the average of the Fisher's Z values was used in the meta-analyses.

Weighting

The weights used in the meta-analyses are based on a random effects model. In this model the relative weights depend on both the sample sizes N_i as used in each original analysis and on the variance of the original effect measures.

The weights are computed as

$weight_i = 1 / var$ (estimate of the effect size T_i)

$= 1 / (var_i(FishersZ) + var(FishersZ_i))$, (cf. Raudenbush, 1994, formula 20-3)

where $var_i (FishersZ) = 1/(N_i-3)$

So the first variance depends on the sample size N_i of a study, the second variance is an overall variance over the estimates of Fisher's Z in all studies used in each meta-analysis.

In order to both reduce the chance factor and to make the computations more simple, $var(FishersZ_i)$ is based on all studies in all analyses and is found to be 0,041. So

$weight_i = 1 / (1/(N_i-3) + 0.041)$.

Fecha de entrada: 28-09-2011 Fecha de aceptación: 01-05-2012