SPRINGER BRIEFS IN EDUCATION

Hans Luyten
Maria Hendriks
Jaap Scheerens

School Size Effects Revisited

A Qualitative and Quantitative Review of the Research Evidence in Primary and Secondary Education

Springer

SpringerBriefs in Education

For further volumes:
http://www.springer.com/series/8914

Hans Luyten · Maria Hendriks
Jaap Scheerens

School Size Effects Revisited

A Qualitative and Quantitative Review of the Research Evidence in Primary and Secondary Education

This study was financially supported by a grant from the PROO (Program Committee for Educational Research), which is part of the Dutch Foundation for Scientific Research (NWO)

Springer

Hans Luyten
Maria Hendriks
Jaap Scheerens
Faculty of Behavioural Sciences,
Department of Research Methodology,
Measurement and Data Analysis
University of Twente
Enschede
The Netherlands

ISSN 2211-1921 ISSN 2211-193X (electronic)
ISBN 978-3-319-06813-8 ISBN 978-3-319-06814-5 (eBook)
DOI 10.1007/978-3-319-06814-5
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014939533

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Contents

Chapter 1
Introduction

Jaap Scheerens, Maria Hendriks and Hans Luyten

Size of school organizations is a recurrent theme in Dutch education policy and has shown fluctuations in the past 20 years. From the mid-1980s until the mid-1990s the government policy has been strongly focused on stimulated scaling-up in all sectors of education, see e.g., the report "Scale and quality in primary education" (Ministerie van Onderwijs and Wetenschappen 1990). The expectation was that scaling-up would be both cost-effective and beneficial to the quality of education and the educational career opportunities for pupils (due to e.g., more choice within larger institutions, easier transfer opportunities to other programs, and more opportunities for professionalization and specialization of staff). From the perspective of school boards, school leaders, and government finally, scaling-up was seen as an important precondition for more decentralization and increased autonomy of schools and institutions. One of the assumptions was that by increasing the autonomy of schools and school boards a more differentiated curriculum would emerge (Onderwijsraad 2005; Ministerie van Onderwijs et al. 2008; NWO 2011; Van de Venne 2006).

Between 1990 and 2006, in all education sectors the number of schools and institutions decreased, while the number of pupils or students within a school or institution increased (Onderwijsraad 2005, 2008). In primary education the average school size in 1990 was 171 pupils, while in and after 2000 an average school had around 220 pupils. In secondary education the mean school size increased from 461 pupils in 1990 to around 1400 in 2006 (Onderwijsraad 2005, 2008; Blank and Haelermans 2008; Ministerie van Onderwijs et al. 2011). It should be noted that in Dutch secondary education schools often comprise several locations. The average number of students per location is approximately 750. Since the turn of the millennium more and more attention is demanded for the side effects and risks of scaling-up. In 2005, the Education Council alerts to the potential risks of ongoing increases in scale, i.e., if these lead to larger educational institutions than is strictly necessary for an effective and efficient performance of their duties. These undesirable effects are related to the freedom of choice of participants and parents, to

H. Luyten et al., *School Size Effects Revisited*, SpringerBriefs in Education,
DOI: 10.1007/978-3-319-06814-5_1,

the management of educational organizations and to the social cohesion within the institutions (Onderwijsraad 2005). Around 2008, a turning point was reached and concern was felt about "the human dimension" in education, as seen from the perspective of pupils, parents, and students (Tweede Kamer 2008). The Ministry of Education, Culture and Sciences in 2008 prepared a memorandum on the human scale in which the human dimension is defined as "an institution being well-organized, so that all those concerned and the stakeholders have a voice and freedom of choice, they all together feel responsible for the school and the lines of decision-making are short" (Tweede Kamer 2008, p. 9). Thereby, scale is seen as an important factor in realizing the human dimension (Ministerie van Onderwijs et al. 2008; Onderwijsraad 2008; Tweede Kamer 2008).

In other countries the same debates with regard to scale are visible (NWO 2011). At the same time, it should be noted there is lack of scientific evidence that underlies the concerns and reforms that are based on it (for the latter, see e.g., the reforms that take place in the US where traditional large high schools are converted into smaller more personal schools, mainly supported by institutions such as the Bill and Melinda Gates Foundation (Kahne et al. 2008; NWO 2011).

In the research on school size effects two main perspectives can be distinguished. On the one hand, there is the basic question of the impact of school size on achievement, which we consider as the effectiveness perspective. On the other hand, research is focused on the cost effectiveness of school size, which is considered the efficiency perspective. A third perspective, which can be seen as a further elaboration of the effectiveness perspective, is the embedding of school size in multilevel school effectiveness models.

1.1 The Effectiveness Perspective: Direct Effects of School Size

What we know from recent review studies and meta-analyses is that both "very small" and "very large" school sizes are less conducive to the quality of education (Cotton 1996; Andrews et al. 2002; Newman et al. 2006; Hendriks et al. 2008; Leithwood and Jantzi 2009). Across studies a different optimum school size is found, which partly seems to be determined by study characteristics such as the country in which the study was conducted and the level of schooling (e.g., primary or secondary education) the study focused on, and the student population characteristics. Another important factor is the type of outcome variable(s) used in the study. In many studies the effect of school size on cognitive outcomes is examined, while other studies focus on social affective outcome measures such as school well-being (see e.g., Stoel 1980), ownership, social cohesion, safety, participation, truancy, drop-out, attitudes toward school or self (see e.g., Andrews et al. 2002; Newman et al. 2006; van de Venne 2006; Feenstra and Gemmeke 2008; Hendriks et al. 2008).

1.2 The Efficiency Perspective

Empirical evidence about the association between school size and costs is limited (van de Venne 2006; Stiefel et al. 2009). Three review studies that pay attention to the economies of school size are available (Andrews et al. 2002; Newman et al. 2006; Leithwood and Jantzi 2009).

1.3 School Size Embedded into Multilevel School Effectiveness Models: Indirect Effects of School Size

In conceptual multilevel school effectiveness models (see e.g. Scheerens 1992; Scheerens and Bosker 1997) school size usually is included as context variable at school level. This implies that school size is more or less perceived as a given condition and not immediately seen as one of the malleable variables that might have a positive impact on achievement. Gaining a better insight into the other preconditions and intermediate school and instruction characteristics that facilitate or impede the effects of school size on outcomes (such as school safety, social cohesion, or participation) is the third perspective of the review study and an important aim of the contractor.

Based on the perspectives three leading questions have been formulated for this review study. The fourth question focuses on school size from the Dutch perspective. The research questions are:

(1) What is the impact of school size on cognitive learning outcomes, noncognitive outcomes and the social distribution of learning outcomes?
(2) What is the "state of the art" of the empirical research on economies of size?
(3) What is the direct and indirect impact of school size, conditioned by other school context variables on student performance? (where indirect effects are perceived as influencing through intermediate school and instruction characteristics)
(4) What is the specific position of the Netherlands in international perspective?

To answer these questions in Chap. 2 an overview is given of the state of affairs of the school size research by focussing on recent review studies on school size effects and Dutch studies that investigated the association between school and different outcome variables. Based on this inventory a tentative conceptual model of school size effects is presented, including different types of preconditions, intermediate variables, and different outcome variables. Next to this, in this chapter the effects of school size on achievement in internationally comparative studies are addressed as well and the results of the scarce Dutch studies that investigated the association between school size and different outcomes.

In Chap. 3 the results of a research synthesis based on the so-called vote count technique are presented combined with a narrative review providing more in-depth

information about school size effects included in the review, thereby focusing explicitly on the correlations with other preconditions and intermediate variables included in the study. The review focuses on a broad set of outcome variables and includes studies that investigated the effects of school size at primary or secondary level of schooling.

Chapter 4 summarizes the effect of school size on student achievement and noncognitive outcomes in a quantitative manner. The approach applied in this chapter yields an overall estimate of expected outcomes at a given school size. Per school size the average standardized outcome across a number of studies are included. The studies included form a subset of the studies covered in the Chap. 2. Separate findings are presented for primary and secondary education.

In the fifth, concluding, chapter the results are summarized and discussed with respect to their relevance for educational policy in general and for the Netherlands in particular. Suggestions for future research on school size are presented.

References

Andrews, M., Duncombe, W., & Yinger, J. (2002). Revisiting economies of size in American education: are we any closer to a consensus? *Economics of Education Review, 21*(3), 245–262.

Blank, J. L. T., & Haelermans, C. M. G. (2008). Trends in onderwijsbureaucratie. *Tijdschrift voor openbare financiën, 2*, 58–73.

Cotton, K. (1996). *School size, school climate, and student performance, close-up #20*. Portland: Northwest Regional Educational Laboratory.

Feenstra, G., & Gemmeke, M. (2008). *De menselijke maat. Effecten van schaalgrootte in het voortgezet onderwijs*. Amsterdam: Regioplan beleidsonderzoek.

Hendriks, M., Scheerens, J., & Steen, R. (2008). *Schaalgrootte en de menselijke maat*. Enschede: Universiteit Twente.

Kahne, J. E., Sporte, S. E., De La Torre, M., & Easton, J. Q. (2008). Small high schools on a larger scale: The impact of school conversions in Chicago. *Educational Evaluation and Policy Analysis, 30*(3), 281–315. doi:10.3102/0162373708319184.

Leithwood, K., & Jantzi, D. (2009). A review of empirical evidence about school size effects: A policy perspective. *Review of Educational Research, 79*(1), 464–490. doi:10.3102/0034654308326158.

Ministerie van Onderwijs en Wetenschappen. (1990). *Schaal en kwaliteit in het basisonderwijs. Projectgroep schaalvergroting basisonderwijs*. Zoetermeer: Ministerie van Onderwijs en Wetenschappen.

Ministerie van Onderwijs, Cultuur en Wetenschap (2008). Onderwijs met ambitie – samen werken aan kwaliteit in het Voortgezet Onderwijs: Kwaliteitsagenda Voortgezet Onderwijs 2008-2011. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschap.

Ministerie van Onderwijs, Cultuur en Wetenschap (2011). *Kerncijfers 2006–2010: Onderwijs, Cultuur en Wetenschap*. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschap.

Newman, M., Garrett, Z., Elbourne, D., Bradley, S., Noden, P., Taylor, J., et al. (2006). Does secondary school size make a difference? A systematic review. *Educational Research Review, 1*(1), 41–60. doi:10.1016/j.edurev.2006.03.001.

NWO. (2011). Programma voor Onderwijsonderzoek (PROO)—Review Studies. Call for proposals 2011. Den Haag: Nederlandse organisatie voor Wetenschappelijk Onderzoek.

Onderwijsraad. (2005). Variëteit in schaal. Keuzevrijheid, sociale samenhang en draagvlak bij grote organisaties. Den Haag: Onderwijsraad.

Onderwijsraad. (2008). De bestuurlijke ontwikkeling van het Nederlandse onderwijs. Den Haag: Onderwijsraad.

Scheerens, J. (1992). *Effective schooling, research, theory and practice*. London: Cassell.

Scheerens, J., & Bosker, R. J. (1997). *The foundations of educational effectiveness*. Oxford: Elsevier Science Ltd.

Stiefel, L., Schwartz, A. E., Iatarola, P., & Chellman, C. C. (2009). Mission matters: the cost of small high schools revisited. *Economics of Education Review, 28*(5), 585–599. doi:10.1016/j.econedurev.2009.01.005.

Tweede kamer, Brief van 28 november 2008 (2008). De menselijke maat in het onderwijs. (BOA/ADV/2008/79252).

van de Venne, L. (2006). Schaalvergroting in het onderwijs. In R. Doorten & R. Rouw (Red.), Raad voor de Maatschappelijke Ontwikkeling. Opbrengsten van sociale investeringen. Amsterdam: Uitgeverij SWP, pp. 105–130.

Chapter 2
School Size Effects: Review and Conceptual Analysis

Jaap Scheerens, Maria Hendriks and Hans Luyten

2.1 Introduction

In this chapter, a review of international review studies on school size effects is presented. Next, ingredients of a more contextualized and tentative causal mediation model of school size effects are discussed. The chapter is completed by a short overview of school size effects as found in international comparative assessment studies and by a synthesis of Dutch empirical studies that have addressed school size effects, in terms of achievement and attainment outcomes, costs, social outcomes, and good teaching practice.

2.2 Review Studies

Early reviews are those by Lee (2000), Cotton (2001), and Andrews et al. (2002).

2.2.1 Lee (2000)

The review study by Lee commented on earlier conceptual literature concerning school size effects. Studies by Conant (1959) and Goodlad (1984) provided no evidence for their recommended optimal number of students for high schools, namely 500. Bryk et al. (1993) found that school size had more influence on social equity than on achievement. This corresponds to other results, to be shown in the sequel, that school size affects students from lower socioeconomic backgrounds more than students from more affluent homes. Still other studies recommended that very large secondary schools ought to be broken up in units of no more than 600, "so that teachers and students can get to know each other" (National

H. Luyten et al., *School Size Effects Revisited*, SpringerBriefs in Education,
DOI: 10.1007/978-3-319-06814-5_2, © The Author(s) 2014

Association of Secondary School Principals 1996, p. 5). Of this and other studies Lee stated that "reformers are out in front of researchers." The tendency from American reviews and conceptual articles went strongly in the direction of "small is beautiful." She then continued in reporting on two empirical studies of her own, in which she did not just look for direct effects of school size on student achievement, but also for indirect effects. From a study in secondary schools Lee and Smith (1997), it was concluded that students learned more in middle-sized secondary schools (600–900), as compared to smaller, but particularly to larger high schools. The size effect was stronger for schools with a large contingent of lower SES students. The overall tendency in a study on 254 elementary schools was that small schools did better, both in terms of direct and indirect effects. The intermediary variable that was used in the elementary school study by Lee and Loeb (2000) was the willingness for teachers to take responsibility for students' learning, and this attitude was more frequently found in smaller schools. Lee offers the following hypotheses for intermediary conditions explaining the effects of school size:

- Less learning in large schools in basic subjects, as a consequence of (perhaps a too broad) offering of curricular options in large schools;
- Teachers less willing to take responsibility for students' learning in large schools;
- More formalized and impersonal social relationships in large schools.

Problems with very small schools could be caused by:

- Disjointed educational experiences, very small schools prone to suffer from just one or two disfunctioning teachers;
- Schools being not large enough to offer a reasonable curriculum.

Lee concluded that there exists no strong research base on school size effects (ibid, 341).

2.2.2 Cotton (2001)

The overview by Cotton (2001) is more like an ideological plea for small schools than a systematic and impartial review of the research evidence. The conclusion is already presented on page 5 of the report: "Research evidence supports decreasing the size of schools to improve student outcomes, school safety and equity, and teacher and parent attitudes." The report only provides references that are supportive of decreasing school size. If one looks at the summary of the report, the long list of favorable assets, ranging from better achievement to "functional

accountability," more inspired staff and better parent and community involvement small schools appear to be the remedy for all educational evils.

Andrews et al. (2002) reviewed the mostly economic literature on economies of size of school districts and schools. Economies of size are defined as the percent change in output resulting from a 1 % increase in all input.

Among the studies on school districts 10 out of 12 empirical studies reported some degree of economies of size. From seven school studies, it appeared that increasing school size from 200 up to 400–600 had little impact on student performance in elementary school. With respect to secondary school studies, Andrews et al. (ibid) repeated the already reported conclusion by Lee and Smith of an optimal size for high schools in the range of 600–900 students. Increasing secondary school size over 1,000 indicates a strong decline in the performance of low SES students.

More fully fledged research reviews are the studies by Newman et al. (2006), Leithwood and Jantzi (2009), and Hendriks et al. (2008).

2.2.3 Newman et al. (2006)

The study by Newman et al. (2006) starts out with listing the most important expectations with respect to the effects of a school being larger or smaller. Large schools are expected to offer wider curricular and extracurricular opportunities, and increased teacher specialization. Smaller schools, on the other hand, are seen as creating a more personalized learning environment, and greater interaction and participation by students and teachers.

Costs are also an important issue as "economies" of scale are expected to occur for larger schools.

Environmental conditions that are associated with the school size issue are the way quasi-market forces impact on schools getting bigger or smaller; with possible implications of discouragement of schools loosing students, and schools as centers of community networks, particularly in rural areas.

In the preparation of his review study, Newman et al. speak of a huge literature, yet from the hundreds of sources that were consulted only 31 studies on secondary schools remained that met basic quality requirements of scientific research. Of these 31 studies 21 originate from the USA and 6 from the UK. Student achievement and attainment was the dependent variable in 19 cases, dropout rate occurred 5 times, student violence 6 times, school climate 5 times, economic outcomes 5 times, and organizational characteristics 2 times. Newman et al. study found the material too diverse and limited to carry out a meta-analyses; the study is described as a narrative review, applying a structured set of rating categories and several raters.

The main conclusions were as follows:

High quality studies, usually focused at student attainment, tended to find quadratic relationships, indicating increase in effects up to a certain size and decline when the schools became still bigger.

Positive effects of school size were more often found when the students were in a higher age category and negative effects were more often found with younger students.

With respect to student behavior and violence as the dependent variables, the results were mixed to a degree that it appeared to be difficult to draw strong conclusions.

Perceptions on school climate appeared to be more positive in smaller schools.

In one study that addressed the relationship between school size and class size, it was found that these are positively related, in bigger schools the average class size tended to be higher.

Finally, the costs per student appeared to be lower in bigger schools.

The author summarized the conclusions by stating that: "The findings of this review would seem to refute some of the more prevalent myths regarding the advantages and disadvantages of smaller and larger school. For example the view that student attainment is universally higher in smaller schools and student behavior is universally worse in larger schools is inconsistent with the current evidence." (ibid, p. 54)

2.2.4 Leithwood and Jantzi (2009)

Leithwood and Jantzi (2009) reviewed 57 "post-1990" studies of school size effects on a variety of student outcomes.

The authors explain the continuing interest in the theme of school size effects by the dynamics of educational systems, where school districts continue to grow or shrink. In their review process, the authors started out with 280 papers of which 57 reported on studies that were selected as useable. Of these 57 studies 40 were targeted at secondary schools and 11 at primary schools, while 6 studies addressed secondary and elementary school size effects. The authors considered the nature of the data reported such that, according to them, a meta-analysis would not be permitted "without eliminating a significant number of studies." So what they did is present a systematic narrative review which included indications on the significance and direction of the results (i.e., associations between school size, and a range of dependent variables). The results are summarized according to the categories of studies used by the authors.

2.2.5 Elementary Schools, Student Achievement

Ten elementary school size effect studies were reviewed. Of these 10 studies 6 reported a negative relationship between school sizes and student achievement (implying that small schools did better), 3 found a nonsignificant relationship, whereas the reader is left wondering what happened to the remaining study.

2.2.6 Secondary Schools, Student Achievement

The number of studies in this category amounted to 19 studies, 15 from the USA, and 3 from the UK. Of these 18 studies, 5 reported a positive relationship, 6 came up with an inverted U shaped distribution of effects, 8 studies found negative associations (small being better). The three studies from the UK were among the ones that reported positive associations. The explanations that were offered for the positive effects (large schools doing better) were "greater opportunities for both instructional and curriculum specialization in larger schools," and the expectation that large schools are likely to have more teachers with specialized skills. Leithwood and Jantzi wondered whether the positive effects of larger schools might perhaps be due to increased dropout rates found in these schools, thus leaving them with a relatively better performing school population. Mentioning possible alternative explanations for positive outcomes (large schools doing better) is one of the instances in this article, where the authors are more critical of studies that find positive effects than of studies which show that small schools do better. Quoting Andrews et al. (2002), the authors say that decreasing returns to size may begin to emerge for high schools above 1,000 students. The authors summarize the results of the studies that had looked into school size effects on student achievement in secondary schools as follows: "While evidence about secondary school size effects on academic achievement is mixed, the most defensible conclusion favors smaller to midsize schools. This conclusion is most accurately portrayed in studies reporting nonlinear relationships between school size and achievement. Lack of attention to dropout rates in studies favoring large schools seriously undermines the confidence we can have in the results." This appears a somewhat partial summary of the evidence, although more negative than positive effects were found. First of all, the fact that studies that established nonlinear relationships do not favor very small schools is neglected in the conclusion. Second, no evidence is reported on the alleged effects of higher dropout in larger schools.

2.2.7 Equity

The authors cite earlier review studies (Lee and Smith 1995; Lee and Loeb 2000; Bickel and Howley 2000) in making the point that many studies find relatively better outcomes for disadvantaged students in smaller schools. Bickel and Howley (ibid) state that "smaller schools would improve schooling in impoverished communities." At the same time, smaller schools do not seem to harm the learning of more advantaged students. Explanatory interpretations on why such outcomes would occur are: the nature of command environments in small schools, less complex subject matter that is learned well, and more attention for disadvantaged students in small schools. Interestingly, the likely fact that smaller schools have smaller classes (due to imperfect matching of number of teachers and number of students in small schools) is not mentioned as a possible explanation.

2.2.8 Attendance, Truancy, and Dropping Out

Leithwood and Jantzi (2009) reviewed 13 studies on school size that used these types of variables as the dependent variable. Two were studies conducted in elementary and 11 were secondary school studies. One studies found a positive effect of school size (large schools doing better), 5 found negative effects (large schools doing worse than small schools on these indicators); 3 studies favored mid-sized schools and 4 studies showed nonsignificant associations. These results are interpreted as strongly favoring small schools. For this category of studies, the authors underline that the studies that found negative relationships were methodologically quite robust. Explanations for the superiority of small schools to foster these kinds of noncognitive outcomes are, firstly that large schools tend to have higher pupil–teacher ratios and that small school advantages are due to how students and teachers relate to each other. "Organizational trust, member commitment to a common purpose, and more frequent contact with people with whom members share their difficulties, uncertainties and ambitions" (Lee and Burkam 2003, p. 385, cited by Leithwood and Jantzi 2009), are considered as assets of small schools.

2.2.9 Participation, Identification, and Commitment to School

For these kinds of outcomes, related to student engagement Leithwood and Jantzi sum up their findings as follows: "Though only 6 studies were located for our review of school size effects on student engagement, they are of quite good quality

and provide entirely consistent evidence in support of the claim that smaller schools are associated with greater student engagement conceived of in several different ways."

2.2.10 Course Taking Patterns

Leithwood and Jantzi (ibid) cite Lee and Smith (1995) who noted that more within school variability in course taking, as is more often the case in large, as compared to small schools, was negatively related to their measures of student performance. Smaller secondary schools show more restrained variability with greater academic emphasis.

2.2.11 Extracurricular Participation

All four studies that Leithwood and Jantzi's (2009) review looked into, indicated that extracurricular participation decreases as school size increases (as a simple linear function). Again these studies are praised for their extremely good quality. The favorable results of small schools are explained by assumed "more pressure on students in smaller schools to participate."

2.2.12 Other Outcomes

Among these other outcomes are student self-esteem, physical safety, and social behavior. The evidence on these outcomes is meager, but always in favor of smaller schools.

2.2.13 Costs and Cost Efficiency

Only five studies looked at these kinds of organizational outcomes of school size, four from the USA and one from Northern Ireland. Two favored large schools, two favored small schools, and one favored midsized schools. Leithwood and Jantzi (ibid) conclude that these five studies offer a clear indication of the most cost-efficient sizes of secondary schools. The authors appear quite interested in the phenomenon of diseconomies of scale, yet, the evidence and explanation on why small schools could be more efficient is quite meager.

2.2.14 Teacher Turnover

On the basis of evidence from two studies, the authors conclude that midsize elementary schools, those in the range of 300 students, may be an optimum size for retaining teachers.

2.2.15 Teacher Attitudes

Out of a total of 10 studies, 7 conducted in elementary schools, and 3 in secondary schools, 7 studies favored smaller schools, 1 study showed a nonlinear relationship and 2 studies showed nonsignificant results. The authors conclude as follows: "While not a unanimous finding, the combined weight of these results seem to indicate that smaller school size enhances the chances that teachers will hold positive work-related attitudes."

The main conclusions that the authors draw are:

"Smaller schools are generally better for most purposes. The weight of evidence provided by the review clearly favors smaller schools for a wide array of student outcomes and most organizational outcomes as well."

"Disadvantaged students are the main benefactors of smaller schools"

"Breath of curriculum is no longer a justification for large schools."

"Cost effectiveness is no longer a justification for large schools, because of higher dropout rates in larger schools."

Next they offer some explanations for the positive balance with respect to smaller schools:

1. A greater sense of community in smaller schools, among students and teachers.
2. A greater sense of identification with the school.
3. More personalized relationships.
4. Teachers knowing their students well.

Next they suggest that school characteristics known from the larger educational effectiveness research literature may be better represented in smaller schools, namely academic press, school disciplinary climate, use of instructional time, teacherś sense of efficacy and teacher quality.

The review ends with clear recommendations to policy makers about optimal school size at elementary and secondary school level, namely 500 and 1,000, respectively. When schools have high proportions of disadvantaged students these numbers should be reduced to 300 for elementary schools and 600 for secondary schools.

Throughout the paper the authors shed doubt on the findings of positive school size effects (large schools having better performance), because the studies in question might have insufficiently controlled for student's dropout. The motivation for this allegation is based on a reference to a study by Rumberger and Palardy

(2005). Leithwood and Jantzi's claim is that larger schools "typically" have larger dropout rates, and that "only few" of the studies that found positive school size effects adequately controlled for dropout. A more precise support of this claim, however, is not provided, and as such it should be seen as a relatively subjective opinion. No figures are presented that compare absolute and relative student's dropout between small and large schools, nor is further information on the supposed selectivity of dropout, namely that it is particularly the low performing students that are dropping out from large schools.

Another one-sidedness in the reporting are repeated remarks about the high quality of the studies that found negative school size effects, while criticizing studies that found positive school size effects. Again these appreciations are not motivated explicitly. The authors do not credibly argue why studies that favor large schools are methodologically weaker.

The review is very limited on the issue of cost-effects. The study by Andrews et al. (2002) is reviewed in a very selective way, underling the occurrence of decreased returns to size, beginning to emerge for high schools above 1,000 students, but omitting the original authors' conclusion that 10 out of the 12 studies that were analyzed found some degree of economies of size.

The paper is also partial in its conclusions when it is stated that positive scale effects are not due to size in itself, but rather to the greater likelihood of more specialized staff. This is a meaningless argument because superiority of small schools, found in other studies, is also explained by making reference to teacher and other intermediary school conditions, associated with size.

One of the more interesting yields of this review is the discussion about the advantages of diversified curriculum offerings, more likely to occur in large schools, as compared to more concentrated curricula associated with smaller schools. Referring back to the article by Lee and Smith (1995), the argument is that the more concentrated curricula have a stronger academic core, which might explain better performance in small American high schools. The fact that in their review of secondary school size effects on student achievement the three UK studies all found positive size effects and the American studies negative school size effects, might be explained by less academically oriented "cafeteria" type curricula in large American high schools, not paralleled in the UK schools. The reader is left wondering what it is about large American high schools that make them less effective than smaller ones, apart from the issue of curriculum emphasis, school composition might be considered as an additional potential explanatory condition, which ideally should be controlled for in size effect studies.

2.2.16 Hendriks, Scheerens, and Steen (2008)

Hendriks et al. (2008) carried out a meta-analysis of the vote-count type, which means that an overview is given of studies that showed significant positive, significant negative or insignificant associations of school size and outcome

Table 2.1 Directions of effects of school size on various dependent variables

	Direction of the effect				Number of	
	−	ns	∩	+	Replications	Publications
Subject						
Achievement	13	23	8	7	51	23
Math	4	7	1	4	16	11
Reading	3	6	1	1	11	8
Science	1	2	0	0	3	3
Other	5	8	6	2	21	15
Social cohesion	12	5	0	2	19	12
Safety	9	9	0	8	26	10
Involvement	10	3	0	1	14	11
Students	8	2	0	0	10	8
Parents	2	1	0	0	3	2
Teachers	0	0	0	1	1	1
Totals*					110	46

− negatively related with school size
ns no significant relation with school size
∩ optimal school size found
+ positively related with school size
*Because publications may refer to more than one dependent variable, the total number of publications is lower than the sum of publications

indicators. Out of a total of 194 originally selected publications, 46 appeared to be useful for this type of analysis.

The results are summarized in Table 2.1.

The results show that the picture of school size effects on student achievement outcomes is quite mixed; a large proportion of the associations is nonsignificant, and about as many of the significant associations are positive as are negative. Results for noncognitive outcomes are different, here negative associations predominate, which means that smaller school size is associated with better results on these indicators.

2.2.17 Conclusions Based on Review Studies

When making up the balance of this review of review studies of the following conclusions can be drawn:

(1) Given the balance in studies that show positive and negative significant associations of school size and cognitive learning outcomes, paired with the large quantity of studies that showed nonsignificant associations, one would be tempted to conclude that school size does not matter for cognitive outcomes.
(2) Further nuance of this conclusion is in place, when one considers the (somewhat limited) number of studies that established nonlinear association,

resulting in estimates of optimal school size. Although these optima are often specified as rather broad intervals, there appears to be a fair degree of consensus on optimal school size ranges for primary and secondary school. The estimates by Leithwood and Jantzi (2009) express this consensus well when they claim that optimal school sizes at elementary and secondary school levels are 500 and 1,000, respectively.

(3) Also, school size seems to matter more for noncognitive outcomes, such as social cohesions, safety, well-being, and involvement.

(4) In the review studies that were analyzed there was a tendency for American studies to show results that favored smaller schools, whereas studies from other countries more frequently found that larger schools did better. Hattie (2009), p. 79 refers to one meta-analysis by Stekelenburg (1991) conducted on the basis of 21 studies on American high schools. The mean effect size was .43, which is substantial. From Hattie's interpretation of these results it becomes clear that the direction of the effect was negative, with smaller schools showing higher student outcomes.

(5) Several reviews confirm the conclusion that school size matters more for disadvantaged than for average students; with disadvantaged students doing better in smaller schools. Leihtwood and Jantzi (2009) propose smaller optimum school sizes, when schools have a large proportion of disadvantaged students (300 for elementary schools and 600 for secondary schools). Most studies establish that smaller size is more important for lower age groups (elementary versus secondary) and early as compared to later grades of secondary schools.

(6) Some reviews are explicitly focused at small schools in rural areas. Social and community outcomes for the school as a center of social activity are counted among the benefits of maintaining small schools in rural communities.

(7) Among the review studies that were analyzed, only two reviews, the ones by Andrews et al. (2002) and Newman et al. (2006) addressed cost issues in relation to school size. Andrews et al. found evidence for economies of scale in 10 out of 12 studies, but suggested that diseconomies of scale may start to occur after the size of secondary schools rises above 1,000 students. Newman et al. found that the cost per students appeared to be generally lower in large schools. A few illustrative studies that we reviewed to follow up these limited results are those by Bickel et al. (2001) and Bowles and Bosworth (2002). Bickel et al. (2002) established that school size has a statistically significant and negative relationship with expenditure per pupil, and noted that cost reduction diminishes as schools grow larger. Bowles and Bosworth (2002), p. 299 summarize the results of their study on seventeen school districts in Wyoming, with the conclusion that: "…averaging across school type, an increase of 10 percent in school size decreases cost per student by approximately 2 percent." In a subsequent section some additional Dutch studies that looked at cost effects of school size will be reviewed. Merkies (2000) used data on Dutch primary schools from the 1986/1987 school year to establish

Preliminary conceptual map

Environment	**School size**	**School organisation; teaching and learning**	**Cognitive and non-cognitive outcomes**
Urban/rural		Class size; school organization; teacher quality; Number of fte management	Reading, mathematics, science; Drop-out rates
SES/ability composition of the school		Climate aspects, discipline, safety	Attitudes towards school, truancy, violence
Age categories of students, elementary and secondary school		Instructional strategies	Attitudes towards learning; academic orientation

Fig. 2.1 A contextualized indirect effect model of school size effects

the relationship between costs and school size. His conclusion is that considerable economies of scale can be acquired by small schools, and that these benefits dissipate as schools grow larger. He also concludes that, from a cost perspective the optimal school size is around 450 pupils and that the costs remain virtually constant from the average school (200 pupils) onwards.

2.3 Toward a Conceptual Model of School Size Effects

School size effects can be studied as direct effects on student outcomes. In that case all likely intermediary variables "between" school size variation and student outcomes are treated as a black box. This is a perfectly legitimate approach, yet, when it comes to explaining school effects, authors cannot do other than to refer to either environmental conditions or intermediary conditions, which are directly affected by changes in school size, and which, in their term may co-vary with educational outcomes. Figure 2.1 sketches a preliminary conceptual map, including types of environmental conditions, types of intermediary variables, and different outcome variables.

2.3.1 Preliminary Conceptual Map

As a next step in this introductory chapter, illustrative research studies will be reviewed that shed light on some of the environmental and intermediary conditions referred to in Fig. 2.1.

2.3.2 Environment

2.3.2.1 Small Schools in Rural Areas

Hargreaves et al. (2009) provide a comprehensive overview of the challenges surrounding small schools in rural areas. On the negative side of small schools are relatively high costs, as well as doubts about the quality of education in small schools. These doubts are persistent but not always supported by the facts, as the authors illustrate in their paper, when referring to urban schools in Scotland, England, and Sweden. They conclude that "there is little research on teaching and learning processes that might account for differential levels of performance, or on how or whether rural schools optimize the resources they have available to them" (ibid, p. 82). On the positive side are all generally accepted advantages of small schools, such as people knowing each other better, more personalized relationships, and easier connections with the local environment. What is also brought into the balance are social benefits for the local community, as the school is sometimes seen as the heart of small rural communities. In a subsequent review of mostly British studies on small rural schools, Hargreaves et al. (2009) provide further details about benefits and challenges of small rural schools. On the positive side they refer to high levels of mutual involvement and collegiality among staff, strong parental involvement and "voice" and the positive esteem for teachers as professionals in rural communities. In the British studies small schools generally came out as being innovative, and show examples of positive effects of multiage classes. Despite this general innovativeness, small schools were somewhat behind in making good benefit of ICT provisions, and showed slow take-up of participation in national headship courses. All in all the review studies by Hargreaves et al. indicate more benefits for small schools than problems. Yet, negative scale effects on costs of very small schools are hard to neglect. Far less convincing is the criticism of lower school organizational and teaching quality in small rural schools. What one might expect is more variance in performance among small schools, as the quality would depend on fewer individuals, offering less opportunity for the leveling out of outlying cases (either very good or very bad teachers) than is the case in larger schools.

2.3.2.2 SES Composition and School Size

Marks (2002) compared the school size issue to the dispute about the effectiveness of catholic versus public schools in the United States. The tendency of American studies is to favor catholic high schools, both with respect to level of achievement outcomes as with respect to equity. Similar results are reviewed with respect to the school size issue (reference to the work of Lee and Smith 1995), emphasizing that small schools reduce the SES achievement relationship, while large schools augment it. Students from disadvantaged and minority background tend to do worse in

large schools. These authors hypothesize measures of social cohesion as intermediary variables between school size and school outcomes, and in this sense they say that small schools may be like catholic schools. A study by Stiefel et al. (2006) found that race-gaps in achievement were negatively related with small schools, in other words, there were smaller race-gaps in smaller schools.

Opdenakker and Van Damme (2007), in a study of Belgium secondary schools, found that school size affects school outcomes positively and that its effect is mediated by school practice characteristics like the amount of collaboration between teachers, which in its turn affected climate and student outcomes. On the basis of structural equation modeling they found that about 25 % of the variance in teacher cooperation could be explained by a joint effect of school size and school composition. Such a joint effect might be interpreted as the effect of school size being "boosted" by school composition (average student ability in the case of this study); and is close to a positive interaction effect of school size and the average ability (or SES) level of the students. In more practical terms; good students tend to do well in large schools.

2.3.2.3 School Size and Age Categories

All studies that compare school size effects between primary and secondary schools (e.g., Lee and Smith 1997; Andrews et al. 2002; Leithwood and Jantzi 2009; Blank et al. 2011) conclude that optimal class size for elementary schools is much smaller than for secondary schools. Blank et al., for example, indicates an optimum size for elementary schools in the range from 440 to 550 students, and for secondary schools 600–1,000. Newman et al. (2006) suggest that in the overall 11–18 age range of secondary schools, the higher age category tolerates larger school size better (ibid, p. 50).

2.3.2.4 Parental Involvement

Involvement of the community with a school, including particularly parental involvement, could be seen as an environmental condition to the school. High community involvement is generally associated with better school performance, although authors like Teddlie et al. (1987) argue that disadvantaged communities would be expected to have a negative influence on school performance. The literature on school size tends to indicate a negative correlation between school size and parental involvement, which implies that small schools tend to have higher parental involvement. This conclusion is confirmed in a study by Dee et al. (2007, cited by Loveless and Hess 2007), who concluded that "the findings provided some tentative evidence that small schools are more effective in promoting parental involvement in schools as well as engagement by the local community." However, they were unable to prove that this conclusion applied to other than rural communities. Walsh (2010) found evidence that the causal direction is from small

schools to parental involvement, and not the other way around, as when involved parents self-select themselves to smaller schools. For parents actually volunteering for certain tasks at schools these self-selection hypotheses could not be rejected. The overall explanation for a decline of parental involvement when schools become larger is the free-rider principle. But, in addition, evidence was found for volunteering parents self-selecting into smaller schools. Finally, Walsh notes that there is evidence that parents see their involvement as a substitute, rather than a complement, for perceived school quality.

2.3.3 School Organization and Teaching/Learning Processes

2.3.3.1 School Size and Class Size

Small classes tend to be clustered in small schools, and average class size is larger in large as compared to small schools (Loveless and Hess 2007). In this way school size effects might "work" indirectly through smaller classes, as intermediary condition. Ready and Lee (2008), cited by Loveless and Hess (2007), found that both smaller schools and smaller classes showed better results in terms of more learning, but the results for small classes were stronger than those for small schools. Part of the reason why small schools may tend to have smaller classes is a sub-optimal match between the number of teachers and the number of classes in smaller schools. Another way to express this would be to say that in smaller schools the optimal or officially allowed class size would be further away from the actual average class size because of the fact that full-time equivalent teachers are undividable. This would provide a clear trade-off between assumed quality enhancement and costs ("involuntary" smaller classes stimulating quality, but raising costs).

2.3.3.2 Bureaucracy and Managerial Overhead

Similar problems of full-time employees being "undividable" would apply to small schools having relatively more managerial overhead than small schools. Blank et al. (2007) established that in Dutch secondary education large schools had relatively less management than small schools. They found no evidence for large schools operating more "bureaucratic" than small schools.

2.3.3.3 Climate Aspects

In the American literature, more personalized relationships and a safe climate are described as some of the major advantages of smaller schools (e.g., Cotton 2001). Such more personalized relationships might be seen as leading to improved

cognitive achievement, but, perhaps more convincingly, to better noncognitive outcomes, in terms of better attendance, less violence and positive attitudes toward school (Newman et al. 2006). Hendriks et al. (2008) found predominantly negative effects of schools size, when social cohesion, safety and school involvement were used as the dependent variable. The Dutch Educational Inspectorate (Inspectie van het Onderwijs 2003) found that students appreciated their (secondary) school better, when it was small as compared to large.

Garrett et al. (2004) on the basis of a review of 31 studies from the USA and the UK found that teachers in smaller schools tended to have more positive perceptions of school climate, of their abilities to influence school policies and school norms, and to control their classrooms; teachers in small school also perceived greater co-operation and resource availability.

Bokdam and Van der Linden (2010) found that secondary school students in the Netherlands found they had better oversight over how their school operated, when the school was below 1000 students in size, and were also more positive over their relationships with teachers.

2.3.3.4 School Size and Curriculum and Instruction

In the American literature, a broader curriculum and more specialized teachers are seen as an advantage of larger schools. At the same time, it may be the case that these broad secondary school curricula, are less academically focused, and offer more opportunities for students opting for a "fun package." The Dutch Inspectorate (2003) reports some differences between strong and weak points of smaller and larger secondary schools in instructional approach. Smaller schools tended to do better in providing structure during lessons and providing clear explanations; small schools did also better in differentiating and providing adaptive instruction. A positive note on larger schools was that, in this study, students thought that larger schools made a better organized impression than smaller schools. In the earlier cited study by Opdenakker and Van Damme (2007), the positive effect of large schools was mediated by better teacher cooperation and classroom climate in larger schools. In the theoretical conjectures put forward by Leithwood and Jantzi (2009) that were cited in a previous section, all kinds of school effectiveness enhancing conditions are associated with smaller schools, but without empirical evidence so far, and meager credibility. Form this preliminary overview of the school size literature, specifying intermediary conditions, the classroom level appears to be a sparsely addressed issue.

2.3.3.5 Conclusion: Partial Evidence on Contextualized Indirect Effect Models of School Size

When venturing a comparison between studies on school leadership effects and school size effects, indirect effect models are even more rarely applied and studied

empirically for school size, than is the case for school leadership (Scheerens 2012). Studying school size effects is simpler on the side of the independent variable specification than studying leadership effects, but school size effectiveness is more complex with respect to the choice of dependent variables and practically unexplored territory as far as intermediary variables are concerned. Next, school size effects appear to depend strongly on modifying conditions, like the age level of students and student background composition, and moreover vary with respect to cognitive and noncognitive outcomes. Finally, the analysis of nonlinear relationships and quadratic functions in school size effects research, is not combined (or combinable) with structural equation modeling of indirect causation models, which is based on the general linear model. The review on potential intermediary conditions in school size effectiveness research has shown very little, in terms of empirical studies actually addressing indirect effect models with the study by Opdenakker and Van Damme (2007) as the only exception.

What remains to be said is to suggest some hypothetical conjectures on plausible variables that might mediate the effect of school size.

2.3.4 Class Size

As a consequence of imperfect matches of full-time teachers to groups of students, average class size is likely to be smaller in small schools. Yet, the degree of class size reduction is not expected to be sizeable, so that the potential explanatory power of this phenomenon is not expected to be strong.

2.3.5 A More Personalized School Climate in Smaller Schools

There is considerable consensus on smaller schools having a more personalized atmosphere with students and teachers knowing each other better. A good relational climate at school is sometimes found to affect cognitive achievement, for example in secondary analyses of the PISA data bases (e.g., Luyten et al. 2005), and there is even stronger evidence that this is also the case for a safe, orderly climate. A more personalized school climate might therefore be a plausible intermediary condition in studies showing better cognitive achievement in smaller schools. Next, this indirect effect would be expected stronger for low SES students and younger students, and be more prominent for noncognitive outcomes such as well-being, involvement and safety.

2.3.6 A More Focused Academic Curriculum in Small Schools

Some of the American studies suggest that small high schools have a more focused academic curriculum than large schools, and that this might be one of the explanations of the often found negative school size effects in the USA. Paradoxically, a more specialized and diversified curriculum is often used as an argument to make schools larger. As suggested earlier, the negative outcomes on large American high schools might be caused by less academic focus in diversified curricular offerings. In the European context, diversified curricula might still be academically focused and this might be a potential explanation for the more frequent positive school size effects in secondary schools.

2.3.7 More Organizational "Modernization" in Larger Schools

In one of the studies that was reviewed, Hargreaves et al. (2009) found that small schools were somewhat slow in picking up ICT applications and leadership courses. Opdenakker and Van Damme (2007) found more teacher cooperation in larger schools. It is not implausible that larger schools invest more in secondary organizational conditions, such as professional development, teacher cooperation, more pronounced and differentiated leadership and technology provisions. To this should be added that the superiority of modernization in secondary organizational processes over more traditional schooling is not a run race, and somewhat more is to be expected of improvements in the primary process of teaching and learning. On this latter issue, association of school size variation and effective teaching, hardly any material was found in the review studies.

2.4 Results from Internationally Comparative Studies

2.4.1 Cross-National Differences in School Size

Data from international comparative assessment studies like PISA and TIMSS show considerable variation in school size between countries. Table 2.2 lists the average school size in 33 OECD countries. The data derive from the PISA 2009 survey,[1] which also includes 40 non-OECD countries. Table 2.2 reports the same figures for these countries. The PISA survey is based on data from 15-year-old

[1] See http://pisa2009.acer.edu.au/downloads.php.

Table 2.2 School size per country in PISA (secondary education)

Country	Mean	Median
OECD (33 countries)		
Australia	761.5	720
Austria	299.0	227
Belgium	553.6	522
Canada	541.9	390
Switzerland	409.0	268
Chile	610.7	470
Czech Republic	362.7	343
Germany	499.0	367
Denmark	403.9	415
Spain	588.8	519
Estonia	412.5	299
Finland	539.1	350
United Kingdom	883.4	869
Greece	205.8	189
Hungary	416.5	337
Ireland	480.7	443
Iceland	274.8	244
Israel	507.3	476
Italy	438.4	343
Japan	500.6	471
Korea	864.4	760
Luxembourg	1104.8	1,022
Mexico	225.5	93
Netherlands	767.3	623
Norway	258.7	238
New Zealand	722.5	583
Poland	297.3	224
Portugal	647.2	603
Slovak Republic	346.8	289
Slovenia	309.5	262
Sweden	430.3	355
Turkey	660.7	474
United States	623.8	366
OECD average (equal weight per country)	**513.6**	
OECD median		**367**
Non-OECD (40 countries)		
Albania	332.9	207
United Arab Emirates	943.9	589
Argentina	327.5	245
Azerbaijan	428.6	319
Bulgaria	398.8	346
Brazil	636.8	507
Colombia	952.9	717
Costa Rica	479.7	313

(continued)

Table 2.2 (continued)

Country	Mean	Median
Georgia	283.0	167
Hong Kong-China	960.7	1,028
Croatia	486.9	457
Indonesia	330.1	201
Jordan	521.9	450
Kazakhstan	405.4	254
Kyrgyzstan	518.8	411
Liechtenstein	192.7	139
Lithuania	398.1	283
Latvia	290.7	181
Macao-China	1318.9	1,359
Republic of Moldova	290.8	222
Malta	488.9	406
Montenegro	738.5	644
Mauritius	657.6	699
Malaysia	1018.9	947
Panama	698.8	476
Peru	272.9	134
Qatar	832.5	571
Shanghai-China	1027.3	851
Himachal Pradesh-India	325.7	298
Tamil Nadu-India	842.9	630
Miranda-Venezuela	565.4	486
Romania	584.7	430
Russian Federation	294.7	188
Singapore	1285.3	1,327
Serbia	619.3	601
Chinese Taipei	1474.2	1,258
Thailand	640.6	375
Trinidad and Tobago	588.9	587
Tunisia	676.5	620
Uruguay	554.1	421
Non-OECD average (equal weight per country)	**617.2**	
Non-OECD median		**440**

students and the findings report the average size of secondary schools. As the distribution of school sizes is rather skewed in most countries (with a bottom effect at the lower end and a long tail at the higher end), the median school size is reported for each country as well. School size is measured as the total number of students enrolled in a school.

On average, the schools in secondary education appear to be somewhat larger in non-OECD countries as compared to OECD countries (617.2 vs. 513.6). In addition, the tables reveal large differences between countries, especially among non-OECD countries. Within the OECD the lowest average school size is reported for Greece

(205.8) and the highest for Luxembourg (1104.8). The average across OECD-countries is 513.6. The average school size in the Netherlands (767.3) clearly exceeds this number. The reported school sizes relate to the numbers of students per location. In the Netherlands, a single school often comprises multiple locations. The average school size for the Netherlands would be about twice as large if entire schools instead of location had been the focus of attention. For only three other OECD countries (Luxembourg, Korea, and the United Kingdom) a mean school size is reported that exceeds the Dutch average. For non-OECD countries the national averages range from 192.7 (Liechtenstein) to 1474.2 (Chinese Taipei).

Table 2.3 lists the per country average school sizes as reported in the TIMSS 2011 survey for 26 OECD countries and Table 2.3 lists the national averages for 24 non-OECD countries.[2] These figures refer to primary schools. On average, the mean school sizes are below those reported for secondary education in PISA. The average school size in primary education across countries in the OECD is 433.5. In the non-OECD countries, it is 744.0. Again, a substantial amount of variation between countries can be observed. The national averages range from 176.8 (Austria) to 1054.0 (Turkey) within the OECD. For non-OECD countries, the range of variation is from 267.4 (Iran) to 1,774 (Qatar).

It should be noted that for OECD countries the average school size in secondary education generally exceeds the size in primary education, whereas this does not apply to non-OECD countries.

The average school size in primary education for the Netherlands (291.3) clearly falls below the cross-national average among OECD-countries. According to figures reported in Dutch sources the number of students per school is even somewhat smaller (Onderwijsraad 2005, 2008; Blank and Haelermans 2008; Ministerie van Onderwijs, Cultuur en Wetenschap 2011). Secondary schools in the Netherlands are relatively large when compared to other OECD countries, but primary schools are particularly small. Only in a small number of other OECD countries (Austria, Germany, Portugal, Ireland, and Northern Ireland) does the average school size in primary education fall below the Dutch average. Figures 2.2 and 2.3 provide a graphical display of the international distribution of school size among OECD countries.

2.4.2 The Effect of School Size on Reading Achievement in PISA

Based on the PISA 2009, data several analyses have been reported that provide information on the relation between school size and reading performance in secondary education across dozens of countries (OECD 2010; pp. 163–188). These

[2] The data are derived from the TIMSS and PIRLS international database, see http://timssandpirls.bc.edu/timss2011/international-database.html.

Table 2.3 School size per country in TIMSS (primary education)

Country	Mean	Median
OECD (26 countries)		
Australia	487.9	433
Austria	176.8	181
Belgium (Flemish)	337.0	311
Chile	740.5	616
Czech Republic	376.0	395
Germany	264.1	246
Denmark	491.0	509
England	333.4	307
Spain	582.5	446
Finland	294.9	283
Hungary	394.9	385
Ireland	279.4	240
Italy	508.7	505
Japan	519.2	528
Korea	1002.0	1,019
Netherlands	291.3	260
Norway	295.9	274
Northern Ireland	288.2	253
New Zealand	357.9	320
Poland	343.4	320
Portugal	219.7	196
Slovak Republic	378.3	356
Slovenia	389.8	383
Sweden	317.9	271
Turkey	1054.0	819
United States	546.4	509
OECD average (equal weight per country)	**433.5**	
OECD median		**338**
Non-OECD (24 countries)		
Armenia	496.4	411
United Arab Emirates	1488.0	854
Azerbaijan	671.0	505
Bahrain	830.5	668
Hong Kong-China	765.4	782
Croatia	607.3	582
Georgia	612.9	491
Iran	267.4	230
Kazakhstan	752.5	650
Kuwait	620.7	609
Lithuania	529.6	450
Malta	378.0	330
Morocco	565.6	516
Oman	548.2	563
Qatar	1774.0	738
Romania	478.8	350

(continued)

Table 2.3 (continued)

Country	Mean	Median
Russian Federation	630.1	616
Saudi Arabia	363.2	314
Singapore	1645.0	1,630
Serbia	730.7	716
Chinese Taipei	1335.0	1,177
Thailand	754.5	333
Tunisia	394.0	353
Yemen	617.7	385
Non-OECD average (equal weight per country)	**744.0**	
Non-OECD median		**539,5**

multilevel regression analyses separately focus on the effect of the five following policy relevant variables:

- School policies on selecting and grouping students
- School governance (e.g., responsibilities for curriculum and assessment)
- School's assessment and accountability policies
- Learning environment (e.g., student–teacher relations, disciplinary climate)
- Resources invested in education (e.g., learning time, class size)

A number of control variables are included in each analysis. These include individual student socioeconomic and demographic background, the school average of the students' economic, social, and cultural status, urban city and school size. The relation between school size and reading performance is modeled as a quadratic function (i.e., both a linear and quadratic term is included in the statistical analysis). In the majority of the per country analyses, the effect of school size is not found to be statistically significant. The average effect across OECD countries is slightly positive in the analyses that focus on the first four variables from the above list. This might point to a somewhat higher level of reading performance in larger schools. However, when controlling for resources, the analysis fails to show an independent effect of school size on average across OECD countries. This seems to imply that across OECD countries reading performance tends to be somewhat higher in larger schools, but that this can be accounted for by the way resources are invested. The following aspects of resource investment are included in the analyses:

- Pre-primary education
- Class size
- Library use
- Extracurricular activities
- Human resources (teacher shortage)
- Quality of educational resources

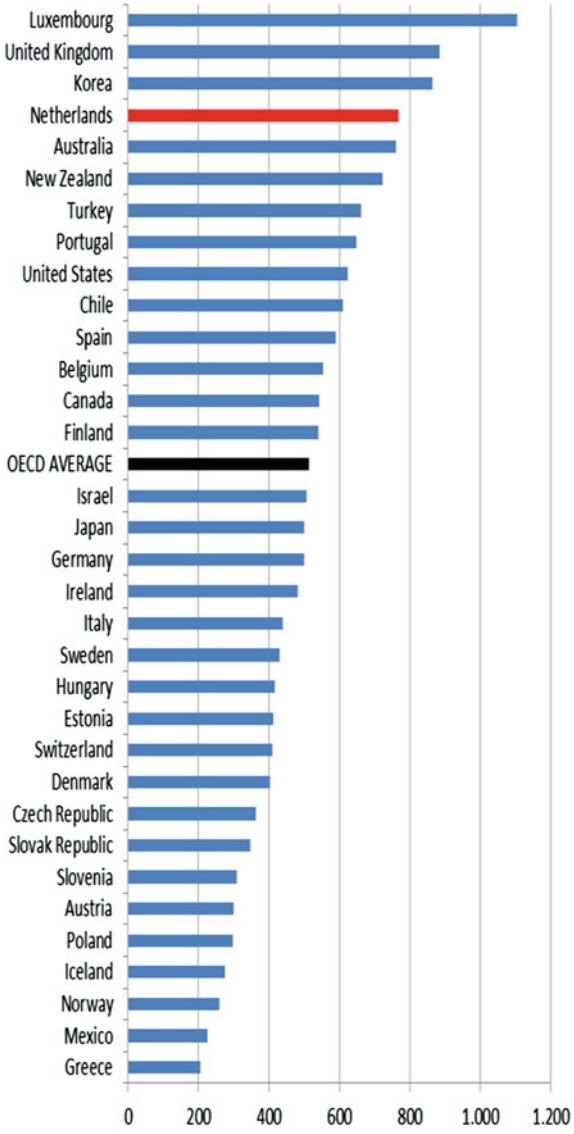

Fig. 2.2 Average school size per country—OECD secondary education. *Source* PISA 2009 dataset

Four OECD countries out of 33 show a significant (and positive) effect of school size in all five analyses (Austria, Belgium, Germany, and Italy). This means that in these four countries the positive trend of higher reading performance in larger schools persists even when student background characteristics, school context, and the aforementioned policy variables are controlled for. In most cases, the effect of the linear term is positive and the effect of the quadratic term is negative. This indicates that the positive effect declines as school size increases.

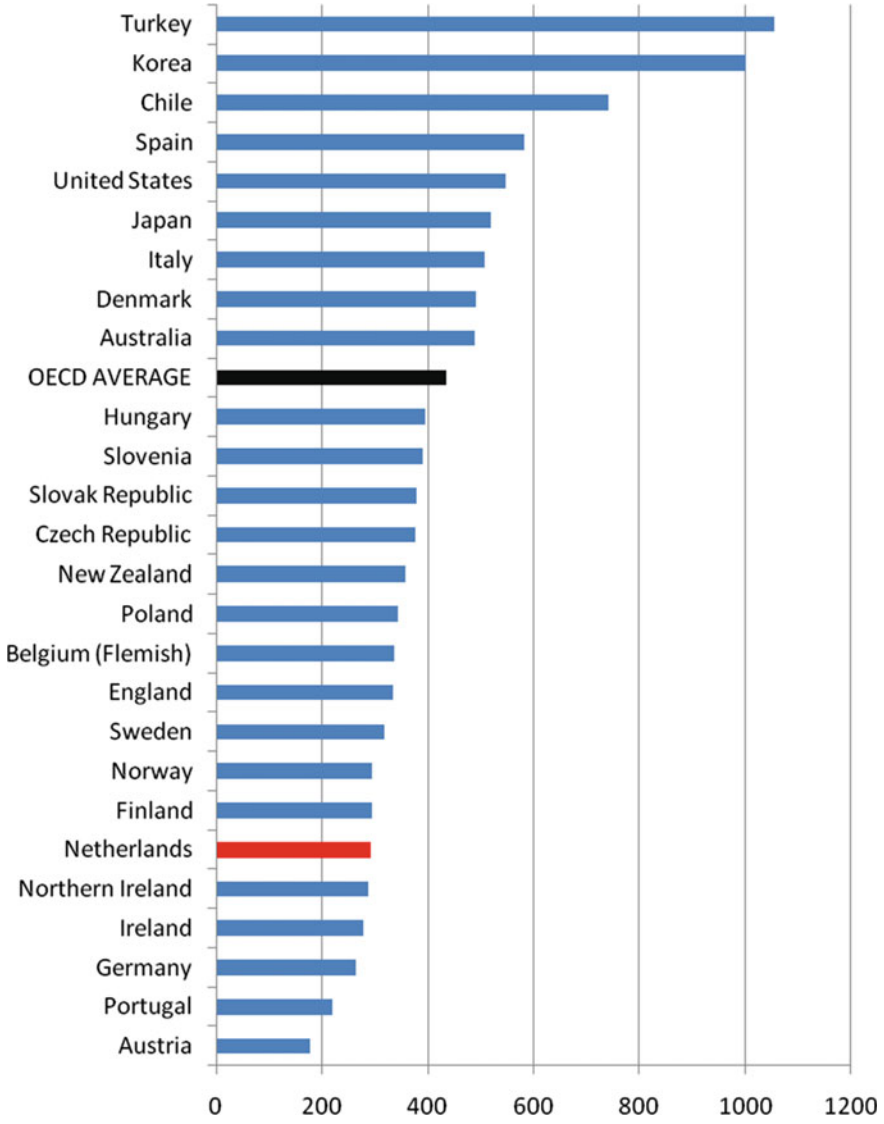

Fig. 2.3 Average school size per country—OECD primary education. *Source* TIMSS 2011 dataset

The effect of school size on reading performance in the Netherlands is found to be stronger than the OECD average in four out of five analyses. However, also in the Dutch case it is found that the effect of school size is reduced to nearly zero (and as a result no longer statistically significant) when taking into account the variables that relate to resources invested.

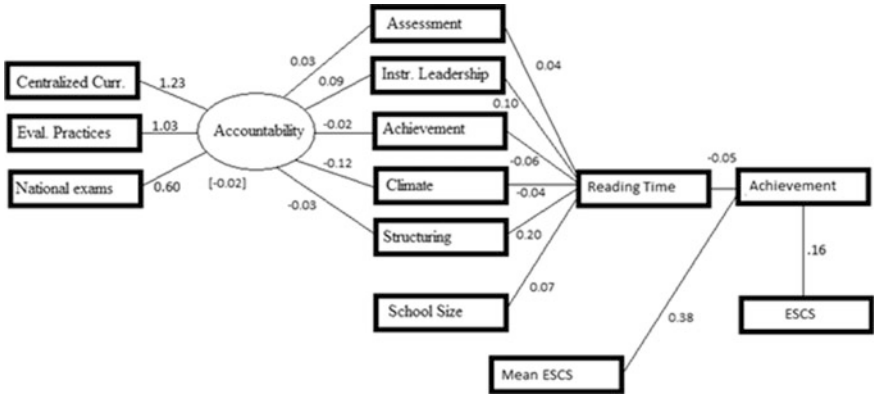

Fig. 2.4 Estimates of the accountability scenario with ESCS. The number in *brackets* is the direct effect of accountability on achievement; cited from Scheerens et al. (2013)

2.4.3 A Closer Look at School Size and Reading Achievement in PISA 2009

As a part of a thematic report on PISA 2009, several multilevel scenarios were analyzed by means of multilevel structural equation modeling (Scheerens et al. 2013). The scenarios included policy amenable variables defined at system level as well as school characteristics and control variables, like student level and school average socioeconomic status of the students. In one of the scenarios, the one that was focused at accountability as the central system level policy amenable condition, school size was included among the school level variables. By way of presenting a schematic overview of the results, the path diagram on the accountability scenario is reprinted as Fig. 2.4. The analyses were conducted on the whole PISA international data base, which, for this scenario had full data on 32 countries.

School size appeared to have a very small positive effect (0.07) on reading time. The indirect effect of time on achievement was negligible. For a more complete discussion of this analysis, the reader is referred to the original report. For the subject at hand, school size appears to have a negligible effect on reading literacy achievement across countries.

2.4.4 Conclusion Based on the Overview of School Size in Internationally Comparative Studies

The descriptive information from the international studies shows considerable variation between country average school sizes. The Netherlands is above the OECD average as far as secondary schools are concerned (744 in the Netherlands

versus 531.6 as the OECD average. With regard to primary schools, the Netherlands is below the OECD average (231.3 as compared to the OECD average of 433.5).

Results from PISA and TIMSS show little relationship between school size and educational achievement. To the extent that a relationship is suggested, this is a small positive rather than a negative effect (better achievement in larger schools). More advanced analyses on the PISA 2009 data set confirmed these results.

2.5 A Closer Look at Dutch Studies on School Size Effects

Of the Dutch studies that were analyzed, three studies looked at achievement outcomes, one study looked at early school leaving as an outcome, three studies investigated cost aspects, six analyzed social outcomes, and one study looked at good teaching practice (the study that was carried out by the Inspectorate.

2.5.1 Achievement and Attainment Outcomes

2.5.1.1 Achievement and School Climate

Dijkgraaf and Van der Geest (2008) and Dijkgraaf and De Jong (2009) used different measures for school size (school district, school, school site, and school track). Using linear and nonlinear models, they found inconsistent effects on student achievement and school climate. Effect sizes were small, often insignificant and if statistically significant, showed a mixed pattern of positive and negative effects. The authors note that if effects were found, this was usually the case when school size was defined in terms of school track (the smallest unit closest to the actual environment where students are taught).

The authors conclude that there is no straightforward, unequivocal relationship between scale and quality in education, which implies that there is insufficient scientific evidence for active educational policy aimed at changing increase or decrease of scale in education.

2.5.1.2 Student Achievement in Math and Science

Luyten (1994) did a study on "School Size Effects on Achievement in Secondary Education" based on evidence from the Netherlands, Sweden, and the USA. The relationship between school size and math and science achievement in the Netherlands was not significant.

2.5.1.3 Achievement (Cito Test Scores)

De Haan et al. (2011), in a study of Dutch primary schools, found that "scale effects can offset the benefits of competition." Changes in the required number of students per schools, decreased competition with 10 %. Contrary to expectations an increase in educational outcomes was found (as, according to economic theory less choice and competition would be expected to lead to a decline in school performance). This outcome was explained by the implication of the policy measure, namely that on average students attended larger schools after the change in required school size, assuming that the decrease in small schools had a positive effect on student performance.

2.5.1.4 Drop Out

Herweijer (2008) starts out by presenting an overview of earlier results that looked into early school leaving (drop) out in secondary educations in relation to school size. The general expectation that larger schools, because of a less personalized atmosphere, would show more early school leaving, is not supported by research.

De Winter (2003) concluded that an optimal size, as far as student well-being is concerned is a school that is neither too big nor too small. Other studies, notably those by Bronneman-Helmers et al. (2002) and by Neuvel (2005) showed no relationship between school size and variables like student well-being and social safety, and studies by van de Venne (2006), and the Educational Inspectorate (Inspectie van het Onderwijs 2003) indicated that there is no relationship between school size and educational achievement of students.

Their own results for Dutch secondary schools show that the bigger the school site is, the smaller the percentage of student that drops out. The authors conclude that their results do not support the supposition that larger schools have more early school leaving.

2.5.2 Costs

Merkies (2000) used data on Dutch primary schools from the 1986/1987 school year to establish the relationship between costs and school size. His conclusion is that considerable economies of scale can be acquired by small schools, and that these benefits dissipate as schools grow larger. He also concludes that, from a cost perspective the optimal school size is around 450 pupils and that the costs remain virtually constant from the average school (200 pupils) onwards.

Blank et al. (2007), in a study of Dutch primary schools, found that the efficiency in terms of the productivity per unit costs of very small schools might be half of that of a larger school. Up to a school size of 300 pupils cost advantages of scaling up occur, after that level cost advantages become gradually smaller, while from a school size that exceeds 550 pupils disadvantages of scale occur.

Blank and Haelermans (2008) document the increase of school size, in all education sectors in the Netherlands in the period between 1990 and 2006. In vocational education, the average school size even became ten times larger.

Changes in school size had implications for the means that are deployed, such as the budget shares for teachers, management, support staff, material costs, and housing. For example, in primary education increased school size led to a larger share of support personnel, but a smaller share of management costs. In secondary education increased school size led to higher cost shares for teachers and material resources, and significantly lower cost shares for management, support staff and housing. In vocational and adult education increase in scale has led to lower cost shares for teachers, support staff and management (taken together) as compared to significant increase in the cost share for material resources and housing.

In all school sectors, except vocational and adult education, economies of scale on expenditure occur. In vocational and adult education, diseconomies of scales were found.

A result of the study, highlighted by the authors, is that despite considerable increase in school size, the share of costs for management and support staff has declined, contrary to expectation of more managerial overhead and bureaucracy in larger schools.

2.5.3 Social Outcomes

2.5.3.1 Student Well-Being

The study by Bokdam and Van der Linden (2010) looked at the way students experienced scale differences in secondary schools. They found that school size is relevant for the degree to which students find their school clearly organized and transparent, and for the quality of the contact with teachers. When school size exceeded 1,000, these two issues appear to suffer, and lead to less quality as perceived by students.

2.5.3.2 Truancy

Bos et al. (1990) found a positive correlation between increase in school size and truancy, implying that truancy becomes more of a problem if school size increases

2.5.3.3 Well-Being and Commitment of Teachers, and Student Teacher Relationships

Feenstra and Gemmeke (2008) carried out a study in Dutch secondary schools in which the relationship between the size of schools or tracks and various facets of

teacher commitment were investigated (commitment of teachers with colleagues and students). The results of their study did not show significant associations of school size with any of these teacher commitment variables.

2.5.3.4 Safety

Mooij et al. (2011) studied multilevel aspects of social cohesion of secondary schools and pupils' feelings of safety. Their main outcome with respect to school size was that students felt safer in larger secondary schools, particularly student who had a background of being bothered by "social violence" (ignoring, excluding, threatening, intimidating, blackmailing, spreading false rumors).

2.5.3.5 Well-Being at School and Safety

van der Vegt et al. (2005) used data from the national school monitor, to study aspects of student well-being and safety in Dutch secondary schools. School size appeared to be not significantly associated with feelings of safety and feelings connected to the school. A significant positive association of school size and safety (bigger schools doing better), was found with respect to the being in place of safety policies. On several other variables bigger schools did worse than smaller schools, namely: more fighting in larger schools, better relationships with peers in smaller schools and more vandalism in larger schools.

2.5.3.6 Teacher Satisfaction

Van der Vegt et al. (2005) studied the effect of school size on teacher satisfaction. They found that the effect of school size was negligible and statistically not significant.

2.5.4 Good Practice in Teaching and Student Attainment

2.5.4.1 Achievement Outcomes, Pedagogical and Didactic Strategies, School Climate, Quality Care at School and Counseling of Individual Students

The Dutch Inspectorate (Inspectie van het Onderwijs 2003) conducted a study about school size and educational quality in 378 secondary schools. The main results were as follows:

– No differences in achievement and attainment results could be attributed to differences in school size.
– Neither could differences in didactic and pedagogical approach, quality care, student support and counseling and school climate be attributed to differences in school size.

2.5.4.2 Conclusions Based on the Dutch Studies

The overview of Dutch studies provides little evidence for scale effects on educational quality, as far as student achievement outcomes, social outcomes (cohesion, well-being, and safety) and even desirable school organizational conditions (teacher satisfaction) are concerned. If significant effects are found, they tend to favor large rather than small schools. See also the review studies by Stoel (1992) and (van de Venne 2006).

With respect to cost efficiency, most authors found a U-shaped development of costs as the size of school increases. Up to a certain level increase in size leads to a decrease in costs, until a certain optimum is reached, beyond this level increase in size leads to increases in average costs.

According to Blank et al. (2011) certain trade-offs can be discerned with respect to scale and quality. The "human measure" may get lost as school size go up, leading to a less personalized school climate. On the other hand, larger schools standardize their production process, by means of tests, quality care systems, and school plans, developed according to standardized formats. Such standardization may have a positive effect on (outcome) quality.

On the basis of these results of Dutch studies, the quality argument might well be put aside in considerations of optimizing school size. Crudely stated: "size does hardly matter for educational quality." Although the gradual trend of cost effects of changes in school size is fairly clear as well, more empirical and analytic work would be useful in the domain of cost effectiveness analyses.

2.6 Overall Conclusions

Review studies show sometimes positive and sometimes negative results. There is a striking difference between US studies as compared to studies in other parts of the world, with studies from the USA indicating more often better outcomes for smaller schools. In terms of expenditure large schools are more efficient, up to a certain threshold.

There is just tentative evidence on the modeling of causal mediation, with school size as the independent and educational outcomes as the dependent variables. Relevant contextual variables that were discussed are: urbanity, SES composition, age category of the school (primary/secondary), and parental

involvement. Variables that might mediate the effect of school size on student performance, which were considered are: class size, managerial overhead, school climate, and facets of curriculum and instruction.

International comparative assessment studies do not show school size as a strong correlate of educational achievement. The very small, usually positive effects (the larger the average school size in a country the better the achievement outcomes) usually disappear when other resources related variables are added in the analyses.

Dutch studies overwhelmingly show that school size does not matter much for educational achievement and social outcomes. The conclusions on costs from Dutch studies are in line with the international state of the art.

References

Andrews, M., Duncombe, W., & Yinger, J. (2002). Revisiting economies of size in American education: are we any closer to a consensus? *Economics of Education Review,21*(3), 245–262.

Bickel, R., & Howley, C. (2000). The influence of scale on school performance: A multi-level extension of the Matthew principle. *Education Policy Analysis Archives,8*(22), 1–33.

Blank, J. L. T., & Haelermans, C. M. G. (2008). Trends in onderwijsbureaucratie. *Tijdschrift voor openbare financiën,2*, 58–73.

Blank, J. L. T., Koot, P. M., & van Hulst, B. L. (2007). *Basisonderwijs en bureaucratie - Een empirisch onderzoek naar de allocatie van middelen in basisscholen*. Delft/Rotterdam: IPSE Studies/ECORYS.

Blank, J., Dumaij, A., & Urlings, T. (2011). *Naar een optimale schaal van publieke voorzieningen*. Delft: IPSE Studies.

Bokdam, J., & van der Linden, B. (2010). *Schoolgrootte uit leerlingperspectief* – Eindrapport. Een onderzoek uitgevoerd in opdracht van het LAKS. Zoetermeer: Research voor beleid.

Bos, K. T., Ruijters, A., & Visscher, A. (1990). Truancy, drop-out, class repeating and their relation with school characteristics. *Educational Research,32*(3), 175–185.

Bowles, T. J., & Bosworth, R. (2002). Scale economies in public education: Evidence from school level data. *Journal of Education Finance,28*, 285–300.

Bronneman-Helmers, H. M., Herweijer, L. J., & Vogels, H. M. G. (2002). *Voortgezet onderwijs in de jaren negentig*. Rijswijk: Sociaal en Cultureel Planbureau, 's-Gravenhage: VUGA.

Bryk, A. S., Lee, V. E., & Holland, R. B. (1993). *Catholic schools and the common good*. Cambridge: Harvard University Press.

Conant, J. B. (1959). *The American high school today*. New York: McGraw-Hill.

Cotton, K. (2001). *New small learning communities: Findings from recent literature* (Vol. 40). Portland: Northwest Regional Educational Laboratory.

Dee, T. S., Ha, W., & Jacob, B. A. (2007). The effects of school size on parental involvement and social capital: evidence from the ELS: 2002. *Brookings Papers on Education Policy* 77–97.

Dijkgraaf, E., & de Jong, M. (2009). Schaaleffecten en kwaliteit. *Economische Statistische Berichten,94*(4553), 87–89.

Dijkgraaf, E., & van der Geest, S. A. (2008). *Schaalgrootte en de kwaliteit van het voortgezet onderwijs*. Rotterdam: SEOR.

Feenstra, G., & Gemmeke, M. (2008). *De menselijke maat. Effecten van schaalgrootte in het voortgezet onderwijs*. Amsterdam: Regioplan beleidsonderzoek.

Garrett Z., Newman M., Elbourne D., Bradley S., Noden P., Taylor J., et al. (2004). Secondary school size: A systematic review. In: *Research evidence in education library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Goodlad, J. (1984). *A place called school: Prospects for the future*. New York: McGraw-Hill.
de Haan, M., Leuven, E., & Oosterbeek, H. (2011). Positieve effecten van schaalvergroting op leerprestaties. *Economische Statistische Berichten,96*(4611), 326–329.
Hargreaves L., Kvalsund, R., & Galton (2009). Reviews of research on rural schools and their communities in British and Nordic countries: Analytical perspectives and cultural meaning. *International Journal of Educational Research, 48*, 80–88.
Hattie, J. (2009). Visible learning. A synthesis of over 800 meta-analyses relating to achievement. London: Routledge.
Hendriks, M., Scheerens, J., & Steen, R. (2008). *Schaalgrootte en de menselijke maat*. Enschede: Universiteit Twente.
Herweijer, L. (2008). *Gestruikeld voor de start, de school verlaten zonder startkwalificatie*. Den Haag: Sociaal Cultureel Planbureau.
Inspectie van het Onderwijs. (2003). *Schoolgrootte en kwaliteit. Groot in kleinschaligheid*. Utrecht: Inspectie van het Onderwijs.
Lee, V. E. (2000). School size and the organization of secondary schools. In M. T. Hallinan (Ed.), *Handbook of the sociology of education* (pp. 327–344). New York: Kluwer Academic/Plenum Publishers.
Lee, V. E., & Burkam, D. T. (2003). Dropping out of high school: The role of school organization and structure. *American Educational Research Journal,40*(2), 353–393.
Lee, V. E., & Loeb, S. (2000). School size in Chicago elementary schools: Effects on teachers' attitudes and students' achievement. *American Educational Research Journal,37*(1), 3–31.
Lee, V. E., & Smith, J. B. (1995). Effects of high school restructuring and size on early gains in achievement and engagement. *Sociology of Education,68*(4), 241–270. doi:10.2307/2112741.
Lee, V. E., & Smith, J. B. (1997). High school size: Which works best, and for whom? *Educational Evaluation and Policy Analysis,19*(3), 205–227.
Leithwood, K., & Jantzi, D. (2009). A review of empirical evidence about school size effects: A policy perspective. *Review of Educational Research,79*(1), 464–490. doi:10.3102/0034654308326158.
Loveless, T., & Hess, F. M. (2007). Introduction: What do we know about school size and class size? *Brookings Papers on Education Policy,2007*(1), 1–14.
Luyten, H. (1994). School size effects on achievement in secondary education: Evidence from the Netherlands, Sweden, and the USA. *School Effectiveness and School Improvement,5*(1), 75–99. doi:10.1080/0924345940050105.
Luyten, J. W., Scheerens, J., Visscher, A. J., Maslowski, R., Witziers, B., & Steen, R. (2005). *School factors related to quality and equity. Results from PISA 2000*. Paris: OECD.
Marks, H. M. (2002). School composition and peer effects in distinctive organizational settings. *International Journal of Educational Research,37*(5), 505–519.
Merkies, A. H. Q. M. (2000). Economics of scale and school consolidation in Dutch Primary School Industry. In J. L. T. Blank (Ed.), *Public provision and performance: Contributions from efficiency and productivity measurement* (pp. 191–218). Amsterdam; New York and Oxford: Elsevier Science, North-Holland.
Ministerie van Onderwijs, Cultuur en Wetenschap. (2011). *Kerncijfers 2006–2010: Onderwijs, Cultuur en Wetenschap*. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschap.
Mooij, T., Smeets, E., & de Wit, W. (2011). Multi-level aspects of social cohesion of secondary schools and pupils' feelings of safety. *British Journal of Educational Psychology,81*(3), 369–390. doi:10.1348/000709910X526614.
National Association of Secondary School Principals. (1996). *Breaking ranks: Changing an American institution*. Reston: Author, in partnership with the Carnegie Foundation for the Advancement of Teaching.
Neuvel, J. (2005). *Monitor sociale veiligheid in de BVE-sector 2004. Deel 1: Deelnemers; Deel 2: Personeel; Deel 3: Beleid*. Den Bosch: Centrum voor Innovatie van Opleidingen.
Newman, M., Garrett, Z., Elbourne, D., Bradley, S., Noden, P., Taylor, J., et al. (2006). Does secondary school size make a difference? A systematic review. *Educational Research Review,1*(1), 41–60.

OECD. (2010). *PISA 2009 Results: What makes a school successful?* (Vol. IV). Paris: OECD.
Onderwijsraad (2005). *Variëteit in schaal. Keuzevrijheid, sociale samenhang en draagvlak bij grote organisaties*. Den Haag: Onderwijsraad.
Onderwijsraad (2008). *De bestuurlijke ontwikkeling van het Nederlandse onderwijs*. Den Haag: Onderwijsraad.
Opdenakker, M.-C., & Van Damme, J. (2007). Do school context, student composition and school leadership affect school practice and outcomes in secondary education? *British Educational Research Journal, 33*(2), 179–206. doi:10.1080/01411920701208233.
Ready, D. D., & Lee, V. E. (2008). Choice, equity, and the schools-within-schools reform. *Teachers College Record,110*(9), 1930–1958.
Rumberger, R. W., & Palardy, G. J. (2005). Test scores, dropout rates, and transfer rates as alternative indicators of high school performance. *American Educational Research Journal, 42*(1), 3–42.
Scheerens, J. (Eds.). (2012). *School leadership effects revisited. Review and Meta-Analysis of Empirical Studies.* Dordrecht: Springer. DOI 10.1007/978-94-007-2768-7.
Scheerens, J., Glas, C., Luyten, H., Jehangir, K., & Steen, R. (2013). *System level correlates of educational performance. Thematic report based on PISA 2009 data*. Enschede: University of Twente.
Stekelenburg, C. R. (1991). *The effects of public high school size on student achievement: a meta-analysis*. Unpublished EdD. University of Georgia, GA.
Stiefel, L., Schwartz, A. E., & Gould Ellen, I. (2006). Disentangling the Racial test Score Gap: Probing the Evidence in a Large Urban School District. *Journal of Policy Analysis and Management*, *26*(1), 7–30.
Stoel, W. G. R. (1992). Schoolgrootte, kosten en kwaliteit: een literatuuronderzoek. In: R. J. Bosker et al. (Ed.), *Schoolgrootte, effectiviteit en de basisvorming*. Enschede: Universiteit Twente, Onderzoekscentrum Toegepaste Onderwijskunde.
Teddlie, C., Stringfield, S., & Wimpelberg, R. (1987). *Contextual differences in effective schooling in Louisiana*. Washington: AERA paper.
van der Vegt, A. L., den Blanken, M., & Hoogeveen, K. (2005). *Nationale scholierenmonitor: meting voorjaar 2005*. Utrecht: Sardes.
van de Venne, L. (2006). Schaalvergroting in het onderwijs. In: R. Doorten, & R. Rouw (Ed.), *Raad voor de Maatschappelijke Ontwikkeling. Opbrengsten van sociale investeringen* (pp. 105–130). Amsterdam: Uitgeverij SWP.
Walsh, P. (2010). Is parental involvement lower at larger schools? *Economics of education review, 29*, 959–970. http://dx.doi.org/10.1016/j.econedurev.2010.04.003
de Winter, M. (2003). *Niet te groot en niet te klein: effecten van schaalgrootte op het welbevinden van jongeren*. Utrecht: NIZW.

Chapter 3
Research Synthesis of Studies Published Between 1990 and 2012

Maria Hendriks

In this chapter, the results of a research synthesis of the effects on school size on various outcome variables are presented. The present review built on an earlier "quick scan" on the impact of secondary school size on achievement, social cohesion, school safety, and involvement conducted for the Dutch Ministry of Education and Sciences in 2008 (Hendriks et al. 2008). It focuses on a broader set of outcome variables, and includes studies that investigated the effects of school size in primary education as well. Studies that provided information about economies of school size were included as well.

The research synthesis seeks answers on the following research questions:

(1) What is the impact of school size on various cognitive and noncognitive outcomes?
(2) What is the "state of the art" of the empirical research on economies of size?

To answer the first question the impact of school size of variety of student, teacher, parents', and school organizational outcome variables was investigated. A distinction is made between outcome variables, i.e., cognitive and noncognitive outcome variables, and school organization variables. Cognitive outcomes refer to student achievement. The noncognitive outcome variables included in the review relate both to students' (attitudes toward school and learning, engagement, attendance, truancy, and drop-out) and teachers outcomes (satisfaction, commitment, and efficacy). School organization variables relate to safety, to involvement of students, teachers and parents, as well as to other aspects of the internal organization of the school, including classroom practices (i.e., aspects of teaching and learning). In the review school organization variables are seen both as a desirable end in itself, but also as intermediate variables conducive to high academic performance and positive student and teacher attitudes. To answer the second question, costs was included as a dependent variable in the review.

In the research synthesis we were not able to apply a quantitative meta-analysis in which effect sizes are combined statistically. One reason was many empirical

H. Luyten et al., *School Size Effects Revisited*, SpringerBriefs in Education,
DOI: 10.1007/978-3-319-06814-5_3, © The Author(s) 2014

studies did not provide sufficient information to permit the calculation of an effect size estimate. What is more, in many of the studies the relationship of school size and a dependent variable is not always modeled as a linear relationship. Instead a log-linear or quadratic relationship is examined or different categories of school size are compared, of which the number and distribution of sizes over categories varies between studies.

Therefore in this research synthesis, we used the so-called vote count technique, which basically consists of counting the number of positive and negative statistically significant and nonsignificant associations. This technique could be seen as a rather primitive form of meta-analysis,[1] and has many limitations, as will be documented in more detail when presenting the analyses. In this chapter, the results of the vote counts with a narrative review providing more in-depth information on a great number of the studies included in the review.

3.1 Search Strategy and Selection Criteria

A computer-assisted literature search procedure was conducted to find empirical studies that investigated the impact of school size on a wide array of student outcomes (such as achievement, cohesion, safety, involvement, participation, attendance, drop-out, and costs). Literature searches of the electronic databases Web of science (www.isiknowledge.com), Scopus (www.scopus.com), ERIC, Psycinfo (provided through Ebscohost), and Picarta were conducted to identify eligible studies. Search terms included key terms used in the meta-analyses by Hendriks et al. (2008), i.e., (a) "school size," "small* schools," "large* schools" (b) effectiveness, achievement (c) cohesion, peer*, climate, community*, "peer relationship," "student teacher relationship" (d) safe*, violence, security (e) influence*, involvement, participation (f) truancy, "drop out," attendance, and (g) costs. In the search, the key terms of the first group were combined with the key terms of each other group separately. We used the limiters publication date January 1990–October 2012 and peer reviewed (ERIC only) to restrict our search.

The initial search in the databases yielded 1,984 references and resulted in 875 unique studies after removing duplicate publications. The titles and abstracts of these publications were screened to determine whether the study met the following criteria:

The study had to include a variable measuring individual school size. Studies investigating schools-within-schools or studies examining size at the school district level were not included in the review. Studies were also excluded if size was measured as grade or cohort enrollment or the number of teachers in the school.

[1] Following Cooper et al. 2009, "vote counting" is still seen as meta-analysis, since it involves statistically describing study outcomes.

The dependent variable of the study had to be one or more of: (1) student attainment and progress, (2) student behavior and attitudes, (3) teacher behavior and attitudes, (4) school organizational practices and teaching and learning, and (5) economic costs

The study had to focus on primary or secondary education (for students aged 6–18). Studies that focused on preschool, kindergarten, or on postsecondary education were excluded.

The study had to be conducted in mainstream education. Studies containing specific samples of students in regular schools (such as students with learning, physical, emotional, or behavioral disabilities) or studies conducted in schools for special education were excluded from the meta-analysis.

The study is published or reported no earlier than January 1990 and before December 2012.

The study had to be written in English, German, or Dutch.

The study had to have estimated in some way the relationship between school size and one or more of the outcome variables. Study had to report original data and outcomes. Existing reviews of the literature were excluded from the review.

When cognitive achievement was the outcome variable studies had to control for a measure of students' background, such as prior cognitive achievement and/or socioeconomic status (SES).

After this first selection, 314 studies left for the full text review phase. In addition recent reviews on school size (i.e., Andrews et al. 2002; Newman et al. 2006; Hendriks et al. 2008; Leithwood and Jantzi 2009) as well as references from the literature review sections from the obtained publication were examined to find additional publications. A cut-off date for obtaining publications was set at 31 December 2012.

The full text review phase resulted in 84 publications covering the period 1990–2012 admitted to the review and fully coded in the coding phase. The data were extracted by one of two reviewers and confirmatory data extraction was carried out by a second reviewer.

3.2 Coding Procedure

Lipsey and Wilson (2001) define two levels at which the data of the study should be coded: the study level and the level of an effect size estimate. The authors define a study as "a set of data collected under a single research plan from a designated sample of respondents" (Lipsey and Wilson, p. 76). A study may contain different samples, when the same research is conducted on different samples of participants (e.g., when students are sampled in different grades, cohorts of students or students in different stages of schooling -primary or secondary-), or when students are sampled in different countries. An estimate is an

effect size, calculated for a quantitative relationship between an independent and dependent variable. As a study may include different measurements of the *independent* variable (school size), as well as different measures of the *dependent* variable (such as e.g., different outcome measures (achievement, engagement, drop-out), different achievement tests covering different domains of subject matter (e.g., language or math), measurement as different point is time (learning gain after 2- and 4 years), a study may yield many effect sizes, each estimate different from the others with regard to some of its details.

The studies selected between 1990 and 2012 were coded by the researchers applying the same coding procedure as used by Scheerens et al. (2007). The coding form included five different sections: report and study identification, characteristics of the independent (school size) variable(s) measured, sample characteristics, study characteristics, and school size effects (effect sizes).

The report and study identification section recorded the author(s), the title and the year of the publication.

The section with characteristics of the explanatory variable(s) measured coded the operational definition of the size variable(s) used in the study (In all studies referring to a measure of total number of students attending a school) as well as the way in which the relationship between size and outcomes was modeled in the study: either linear or transformed to its logarithm (size measured as a continuous variable), quadratic (estimating both linear and quadratic coefficients), or comparing different size categories.

The sample characteristics section recorded the study setting and participants. For study setting the country or countries in which the study was conducted were corded. With regard to participants, the stage of schooling (primary or secondary level) the sample referred to was coded as well as the grade or age level(s) of the students the sample focused on. The number of schools, classes, and students included in the sample were recorded as well.

The study characteristics section coded the research design chosen, the type of instruments employed to measure the time variable(s), the statistical techniques conducted and the model specification. For the type of research design, we coded whether the study applied a quasi experimental—or experimental research design and whether or not a correlational survey design was used. With regard to the type of instruments used we coded whether a survey instrument or log was used and who the respondents were (students, teachers, principals, and/or students), and whether data were collected by means of classroom observation or video-analysis or (quasi-) experimental manipulation. The studies were further categorized according to the statistical techniques conducted to investigate the association between time and achievement. The following main categories were employed: ANOVA, Pearson correlation analysis, regression analysis, path analysis/LISREL/SEM, and multilevel analysis. We also coded whether the study accounted for covariates at the student level, i.e., if the study controlled for prior achievement, ability, and/or student social background.

Finally, the school size effects section recorded the effects sizes, either taken directly from the selected publications or calculated. The effect sizes were coded

as reflecting the types of outcome variables distinguished in the review (i.e., achievement, students' and teachers' attitudes to school, students', teachers' and parents' participation, safety, attendance, absenteeism, truancy and drop out, school organization and teaching and learning, and costs). With regard to achievement, four groups of academic subjects were distinguished in the coding: language, mathematics, science, and other subjects.

3.3 Vote Counting Procedure

As the nature of the data reported in the 84 studies and 107 samples did not permit a quantitative meta-analysis without eliminating a significant number of studies in each of the outcome domains, a vote counting procedure was applied. Vote counting permitted inclusion of those studies and samples that reported on the significance and direction of the association of school size and an outcome measure, but did not provide sufficient information to permit the calculation of an effect size estimate. Vote counting comes down to counting the number of positive significant, negative significant, and nonsignificant associations between an independent variable and a specific dependent variable of interest from a given set of studies at a specified significance level, in this case school size and different outcome measures (Bushman and Wang 2009). We used a significance level of $\alpha = 0.05$. When multiple effect size estimates were reported in a study, each effect was individually included in the vote counts. Vote counting procedures were applied for each of the (groups of) dependent variables: achievement, students' and teachers' attitudes to school, students', teachers' and parents' participation, safety, attendance, absenteeism, truancy and drop out, school organization and teaching and learning, and costs.

The vote counting procedure has been criticized on several grounds (Borenstein et al. 2009; Bushman 1994; Bushman and Wang 2009; Scheerens et al. 2005). It does not incorporate sample size into the vote. As sample sizes increase, the probability of obtaining statistically significant results increase. Next, the vote counting procedure does not allow the researcher to determine which treatment is the best in an absolute sense as it does not provide an effect size estimate. Finally, when multiple effects are reported in a study, such a study has a larger influence on the results of the vote-count procedure than a study where only one effect is reported.

As vote counting is less powerful it should not be seen as a full blown alternative to the quantitative synthesis of effect sizes, but, rather as a complementary strategy.

Table 3.1 gives an overview of the studies, samples, and estimates included in the vote counting procedures for each type of outcome variables (i.e., achievement, students' and teachers' attitudes to school, students', teachers' and parents' participation, safety, attendance, absenteeism, truancy and drop out, school organization and teaching and learning, and costs) as well as in total.

Table 3.1 Number of studies, samples and estimates included in the vote counting procedure for each (group of) dependent variable(s) and in total

	Studies	Samples	Effect size estimates
Achievement	46	64	126
Students' and teachers' attitudes to school	14	14	24
Participation	10	13	13
Safety	24	25	54
Attendance, absenteeism, and truancy	12	16	23
Drop-out	4	5	5
Other student outcomes	5	7	9
School organization and teaching and learning	4	4	18
Costs	5	5	5
Total	84	107	277

3.4 Moderator Analysis

Moderator analyses were conducted to examine the degree to which the relationship between school size on one hand and an outcome variable on the other could be attributed to specific sample or study characteristics. Due to the low number of samples included in the review for most of the outcome variables (see Table 3.1), moderator analysis was only applied for those studies and samples that included student achievement or safety as the outcome variable, and in which the relationship between size and outcomes was modeled as a linear or log-linear function. The following types of moderator variables were used in our analyses: sample characteristics as geographical region, and the level of schooling (primary, secondary schools), and study characteristics that refer to methodological and statistical aspects, e.g., study design, model specification, whether or not covariates at the student level (SES, cognitive aptitude, prior achievement) or school level (school level SES, urbanicity) are taken into account and whether or not multilevel analysis was employed.

3.4.1 Characteristics of the Studies and Samples Included in the Review

In total, 84 studies and 107 samples were included in the review. Almost three quarter of the studies (i.e., 58 studies) originate from the United States. Seven studies were conducted in the Netherlands, four in the United Kingdom, three in Israel, two in Canada, two in Sweden, and one in each of Australia, Hong Kong, Ireland, Italy, and Taiwan.

Eighteen studies examined effects of school size in primary education contexts, 53 studies in secondary schools, and six studies collected data in primary and

secondary schools separately. In three studies, a combined sample of primary and secondary schools was used.

More detailed information about the characteristics of the samples and studies can be found in Tables A.1 and A.2.

3.5 Results

3.5.1 Academic Achievement

Evidence about the relationship between school size and academic achievement was derived from 46 studies and 64 samples (yielding in total 126 effect estimates). Of the 46 studies, 20 studies (22 samples) provided evidence about the relationship between school size and achievement in primary education. Evidence about the effects of school size in secondary education was available from 29 studies (39 samples). In five studies the data were obtained from samples that included students from both levels of schooling. The majority of studies (and samples) were conducted in the United States. The other studies originate from Canada (1 sample), Hong Kong (1 sample), The Netherlands (2 samples), and Sweden (2 samples).

Table 3.2 shows the results of the total number of negative, nonsignificant, curvilinear, and positive effects found for the associations between school size and achievement. In this table, evidence is presented for all studies in total as well as separately for the three different ways in which school size is measured in the studies: (1) school size measured as a continuous variable usually operationalized as the total number of students attending a school or different sites of a school at a given date, suggesting a linear relationship, (2) school size measured as a quadratic function, seeking evidence for a curvilinear relationship and, (3) school size measured through comparison of different categories. In these latter studies, the evidence reported could show either a linear or curvilinear relationship, or favoring a certain size category.

The results of the vote counting show that of 126 effects sizes in total, more than half of the associations (78 effects, 62 %) between school size and achievement appeared to be nonsignificant, 23 estimates (18 %) showed negative effects and 11 estimates (9 %) positive effects.

3.5.2 School Size Measured as a Continuous Variable

When school size was measured as a continuous variable, in 11 of the 46 samples (20 effects) a negative relationship between school size and achievement was reported while in 8 samples (8 effect sizes) it was found that achievement rises as school size increases (see Tables 3.2 and A3).

Table 3.2 Results of vote counts examining the number of negative, nonsignificant, curvilinear, and positive effects of school size on achievement

	Studies	Samples	Direction of effect			
			−	ns	∩	+
School size measured as a continuous variable	31	46	20	62	0	8
School size squared measured	4	8	0	0	8	0
School size measured as discrete variable (categories)	15	18	3	16	6	3
Total	46	64	23	78	14	11

− = negatively related with school size
ns = no significant relation with school size
∩ = optimal school size found
+ = positively related with school size

In 15 of the 46 samples effects were examined for more than one achievement measure (e.g. in different domains (language or math), or at different points in time), the effects reported within one sample were in the same direction, thus all effects found were either nonsignificant, positive, or negative. The only sample that reported conflicting results was the study by Fowler and Walberg (1991). In this study 13 school achievement outcome measures were regressed on 23 school characteristics. After district socioeconomic status and the percentage of students from low-income families were accounted for, school size was the next most influential and consistent factor related to outcomes. Five of the achievement measures were negatively associated with school size; the other effects were nonsignificant. According to the authors these results suggest "that smaller school districts and smaller schools, regardless of socioeconomic status …, may be more efficient at enhancing educational outcomes" (p. 189). However, other authors (Spielhofer et al. 2004) recommended caution as only school level data were used in the regression analysis.

Besides Fowler and Walberg, eight other studies (samples) also found negative associations between school size and achievement (Archibald 2006; Caldas 1993; Deller and Rudnicki 1993; Driscoll et al. 2003; Heck 1993; Lee and Smith 1995; Moe 2009; Stiefel et al. 2006). In four of these studies the effect of school size on achievement was examined at different levels of schooling (Caldas 1993; Driscoll et al. 2003; Moe 2009; Stiefel et al. 2006). In these four studies the authors all reported a (weak) negative effect for primary education while for secondary education a nonsignificant (negative or positive) effect was found. Two of the remaining studies were conducted in primary education (Archibald 2006; Deller and Rudnicki 1993) and in the study by Heck a sample from both primary and secondary schools was used.

Archibald conducted the study in Washoe County, Neveda, USA. The researcher used a three level HLM model and found a small negative relationship

between school size and both math and reading (standardized regression coefficient $\beta = -0.03$ for reading and -0.07 for math).

Ma and McIntyre examined the effects of pure and applied mathematics courses on math achievement in Canada, using data from the Longitudinal Study of Mathematics Participation. Variables included in the multilevel model were student background variables, prior math achievement, course attendance (pure math, applied math, low-level preparatory math), school location, school SES, parental involvement, and school climate. Ma and McIntyre did not find a significant main effect. In the final model positive interaction effects of school size with course taking were found. Students taking pure math or students taking applied math in smaller schools had higher achievement in math than did students taking pure math or applied math in larger schools. The effects were small: "a difference of 100 students in enrollment was associated with a difference in mathematics achievement of 5 % of a standard deviation. A quarter of a standard deviation often indicates a difference that is substantial enough to warrant practical implications: to reach that level a reduction in school size between 400 and 500 students is required" (p. 843).

Five studies (8 samples) found positive effects, i.e., achievement declined as school size increased (Borland and Howsen 2003; Bradley and Taylor 1998; Foreman-Peck and Foreman-Peck 2006; Lubienski et al. 2008; Sun et al. 2012). For three of these studies the curvilinear relationship was examined as well (for these studies see the text on curvilinear relations below).

In the study by Lubienski et al. (2008) the relationship between school size and math achievement is examined both in primary and secondary education, using data from the National Assessment of Educational Progress (NAEP 2003) on over 150,000 students in grade 4 (primary) and 110,000 in grade 8 (secondary). Variables included in the HLM models refer to school type, student demographics, school demographics, school location, school climate, teacher education and experience, teaching methods, and student beliefs and attitudes. The authors found that "school size is slightly positive associated with math achievement" (p. 129) in grade 8, and nonsignificant in grade 4. Moreover, they noted that the "demographic variables accounted for the vast majority of the variance in achievement between schools" (p. 128).

In the study by Sun et al. (2012) data were taken from the Hong Kong sample of PISA 2006. The dependent variable was science literacy. For statistical analysis, the authors used a two-level multilevel model. At the student level sex (male students performed better), student SES, parental views on science, motivation, and student self-efficacy positively contributed to student science achievement. At the school level, school SES composition, quantity of instruction and school size were found to be positive predictors of science achievement. A possible explanation the authors provide for the positive effect of school size on science achievement is that "larger student body schools are more likely to have more grants or financial opportunities and greater support from parents … Therefore, big schools are more likely to attract and retain qualified and talented science teachers as well as create large peer effects as more active and bright students work together" (p. 2118).

3.5.3 Curvilinear Relationships (School Size as a Quadratic Function)

Of the 46 samples in which school continuous variable, 8 samples (4 studies) also reported curvilinear relationships (Borland and Howsen 2003; Bradley and Taylor 1998; Foreman-Peck and Foreman-Peck 2006; Sawkins 2002) (see Tables 3.3 and A.4).

The study of Borland and Howsen is the only study providing evidence about the curvilinear relationship of school size effects on academic achievement of elementary (3rd grade) students. The study was conducted in Kentucky (United States). The mean school size of the 654 schools was 490 students. Other variables in the model included student ability, teacher experience, the existence of a teacher union, average income of the community, class size, and poverty. The results of the two-stage least-squares regression suggested an optimal school size of around 760 students.

The three studies related to secondary education were all conducted in the United Kingdom. All three studies focused upon the upper end of the exam results distribution, with either the proportion of 15–16-year-old pupils in each school obtaining five or more General Certificate of Secondary Education (GCSE) examination results at grades A to C in England (Bradley and Taylor) or Wales (Foreman-Peck and Foreman-Peck) as dependent variable, or the percentage of pupils in their last year of secondary education (S4) gaining five or more Standard Grade passes at levels 1 or 2 in Scotland (Sawkins). The estimates for the samples in England and Wales suggest an inverted 'U'-shaped relationship between school examination performance and school size. For the schools in England (Bradley and Taylor) the optimum school size found was around 1,200 students for 11–16 schools and 1,500 students for 11–18 schools, optima that seem to be considerably higher than the mean school size of the schools in the samples (685–765 for 11–16 schools and 916–1,010 for 11–18 schools, see also Table 3.3). The optimum school size found for schools in Wales appeared to be much lower (560 students), both compared to the evidence in England and to the mean sizes of the schools in the Welsh samples (respectively 871 in 1996 and 936 in 2002).

In the study using Scottish data (Sawkins 2002), a contradictory 'U'-shaped relationship was found between examination performance and school size. Scottish school examination performance appeared to decline as the number of pupils in a school increases, reaching a minimum turning point of around 1,190 pupils for the 1993–1994 sample and 1,230 pupils for the 1998–1999 sample, after which the performance started to increase. The explanation might be that in Scotland very large schools are uncommon. In the study by Sawkins, only 4 % of the secondary schools appeared to be larger than the calculated minimum.

Table 3.3 Overview of directions of effect (negative, nonsignificant, positive, and curvilinear) of relationships of school size on academic achievement for each sample (school size effect modeled as quadratic function)

Study	Sample	School level	Direction of effect				Remarks
			−	*ns*	+	∩	
Borland and Howsen (2003)		P				∩ 760	Linear (+)
Bradley and Taylor (1998)	11–16 1992	S				∩ 1130	Linear (+)
	11–16 1996	S				∩ 1230	Linear (+)
	11–18 1992	S				∩ 1350	Linear (+)
	11–18 1996	S				∩ 1440	Linear (+)
Foreman-Peck and Foreman-Peck (2006)		S				∩ 560	Linear (+)
Sawkins (2002)	1993–1994	S				U 1190	Linear (−) Only 4 % of schools were larger than the calculated minimum for 1993–1994
	1998–1999	S				U 1230	Linear (−) Only 3.3 % of schools were larger than the calculated minimum for 1998–1999

P = primary, *S* = Secondary
− = negatively related with school size
ns = no significant relation with school size
∩ = optimal school size found
+ = positively related with school size

3.5.4 School Size Measured as Categories

In 15 studies (18 samples), schools were classified in categories, based on the number of pupils. Six studies (6 samples), were conducted in primary education and 10 studies (8 samples) in secondary education (see Table A.5). The range of school sizes included in the studies was variable. Some studies compared small and larger schools while in other studies schools of three or more different size categories were compared.

In three samples (2 studies), a positive relationship between school size and achievement was found (large schools doing better) (Gardner et al. 2000; McMillen 2004) and in three other samples (2 studies) a negative association (Eberts et al. 1990; Lee and Loeb 2000). In 16 samples, the relationship was nonsignificant, and in the remaining six samples a certain size category or optimum was favored (Alspaugh 2004; Lee and Smith 1997; Ready and Lee 2006; Rumberger and Palardy 2005).

In their study of 264 inner-city elementary schools in Chicago Lee and Loeb (2000) found that school size influenced both teachers and students. In small schools (with 400 pupils or less) 1 year gains in math achievement were significantly higher compared to those in mid-size (400–750 pupils) and large schools 750 pupils or more). Both direct and (small) indirect effects were found, the latter through teachers' positive attitudes about collective responsibility for student learning. The limited number of small schools participating in the study, however, was a drawback of the study. Only 25 of the 264 schools were small (400 pupils or less).

McMillen (2004) investigated the impact of school size achievement for three separate samples of students (at either elementary, middle, or high school level), using longitudinal achievement data from schools in North Carolina. At high school level, a positive and main effect was found of school size with both reading and math achievement after controlling for school and student demographic characteristics. Students in larger high schools were associated with higher achievement. But "the benefits of size at the high school level, however, appeared to accrue disproportionally … to higher–achieving students, white students and students whose parents had more education, especially in mathematics …" (p. 18). At the elementary and middle cohort the multilevel analyses yielded no statistical significant main effects for school size, but small interaction effects were found between size and prior achievement. Students who scored on grade level in the 3rd (respectively 6th) grade tended to do slightly better in larger middle and high schools. Students who scored below grade level in grade 3 (respectively 6) performed better in smaller schools. The interaction effects found at high school level (between size and ethnicity and size and parent education) were nonsignificant at primary and middle school level. McMillen also estimated curvilinear effects for school size. However, in all models tested, a better fit was achieved when only the linear term for school size was used. Possible explanations for the results found in the study refer to the broader curriculum offerings in large schools (higher

achieving students in large schools might be able to take more advantage of these) (see also Haller et al. 1990; Monk 1994), and/or the culture and organization of small schools. Students from disadvantaged and minority background might have better achievement in small schools because of the better social climate and more personal relationships between students and teachers.

Rumberger and Palardy (2005) used data from the National Education Longitudinal Study (Nels: 88) to estimate the impact of school size on achievement growth, drop-out rate and transfer rate. The study was based on a sample of 14,199 pupils from 912 schools in the United States (nationwide) and was one of the rare studies in which achievement growth and drop-out rate were investigated simultaneously. Results of the multilevel analyses showed that "schools that are effective in promoting student learning (growth in achievement) are not necessarily effective in reducing drop-out and transfer rates" (p. 24). An "inverted U" relationship was found for achievement and drop-out. Achievement growth was significantly higher in large high schools (1200–1800 pupils) as was also the drop-out rate. Next to this, it was found that background characteristics contributed differently to the variability in the various outcome measures (i.e., 58 % of the variance in school drop-out rates, 36 % of the variance in student achievement and 3 % of the variance in transfer) as did also school policies and practices. When dropout was the dependent variable, school policies and practices accounted for 25 % of the remaining variance after controlling for student background. This was far more than for achievement or transfer.

The study by Luyten (1994) is the only Dutch study examining the association between school size and achievement included in the review. Luyten employed multilevel analysis to investigate the effect of school size on math and science achievement in the Netherlands, Sweden, and the US. Controlling for background characteristics (sex, achievement motivation, socioeconomic status, and cognitive aptitude), the study did not reveal any significant effects in any of the three countries.

3.5.5 Moderator Analyses

For the studies and samples in which school size was measured as a continuous variable moderator analyses were conducted to examine the degree to which the relationship between school size and achievement could be attributed to specific characteristics of the study or sample. Also we investigated whether the school size and achievement correlation was moderated by the academic subjects in the achievement measure.

The analyses of vote counts applied to studies and samples addressing the impact of school size on achievement in different subject areas does not show differences of importance (see Table 3.4). The percentage of positive effects (students in larger schools having better performance) for achievement in science and "all other subjects" is somewhat higher than those for language and mathematics.

Table 3.4 Results of vote counts examining the number and percentage of negative, nonsignificant, and positive effects of school size on academic achievement in all subjects, language, mathematics, science, and subjects other than math or language (school size measured as a continuous variable)

Subject	Negative effects N	Nonsignificant effects N	Positive effects N	Negative effects %	Nonsignificant effects %	Positive effects %
All subjects	20	62	8	22	69	9
Subject math	5	19	1	20	76	4
Subject language	7	19	0	26	74	0
Subject science	1	4	1	17	67	17
Subject other than math or language	7	20	6	21	61	18

Moderator analyses of study and sample characteristics examining the number and percentage of negative, nonsignificant, and positive effects of school size on academic achievement are presented in Table 3.5. Of the moderator analyses of study and sample characteristics, the statistical technique employed and the inclusion of a covariate for student's prior achievement in the model tested are the most striking outcomes. More negative effects are found in studies that account for prior achievement as well as in studies that employed multilevel modeling.

3.5.6 Social Cohesion: Attitudes of Students and Teachers Toward School

Fourteen studies (15 samples, yielding in total 26 effect estimates) provided evidence about the relationship between school size and students' and teacher attitudes toward school (see Tables 3.7, A.6, A.7). Evidence about the effects of school size on attitudes was mainly available from secondary education (12 studies; 13 samples). Only two of the 14 studies examined the impact of school size on students' attitudes in primary education.

The majority of studies were conducted in the United States (9 studies; 10 samples). Other countries were Australia (1 study), Israel (1 study), Italy (1 study), and the Netherlands (2 studies).

The outcome variables (attitudes) measured in the studies could be classified into three main variables: identification and connection to school, relationships with students, and relationships with teachers (see Table 3.6). With regard to student attitudes identification and connectedness to schools the variables used included perceptions of pupils' like feeling part of the school, feeling competent and motivated, feeling safe, being happy and satisfied with school, with education and the usefulness of their school work in later life. Relationships with students targeted at perceptions of being happy together as well as the kindness and

Table 3.5 Results of moderator analyses examining the number and percentage of negative, nonsignificant and positive effects of school size on academic achievement (school size measured as continuous variable)

Moderator	Negative effects N	Nonsignificant effects N	Positive effects N	Negative effects %	Nonsignificant effects %	Positive effects %
Level of schooling						
Primary school	7	24	1	22	75	3
Primary and secondary school	2	3	0	40	60	0
Secondary school	11	35	7	21	66	13
Country						
Canada	0	1	0	0	100	0
Hong Kong	0	0	1	0	0	100
Netherlands	0	2	0	0	100	0
Sweden	0	1	0	0	100	0
UK	2	5	5	17	42	42
USA	18	53	2	25	73	3
Covariates included						
Included covariate for student's prior achievement	8	15	1	33	63	4
Included covariate for ability	0	3	1	0	75	25
Included covariate for SES	8	23	3	24	68	9
Included covariate for composite SES	19	57	8	23	68	11
Included covariate for urbanicity	2	5	1	25	63	13
Statistical technique used						
Technique multilevel	7	13	2	32	59	9
Technique not multilevel	13	49	6	19	72	0
Total	20	62	8	22	69	9

helpfulness of their peers. The relationship with teachers is a variable in which relational aspects were included (e.g., the teacher treats pupils fairly and cares about them) as well as perceptions with regard to the support students receive (such as encouraging students to higher academic performance, helping pupils with school work).

As identification and connection to school is concerned, Kirkpatrick Johnson et al. (2001) distinguish between affective aspects (the feelings toward and identification with school, which he calls school attachment) and behavioral aspects

Table 3.6 Overview of outcome variables and variable heading used in studies where attitudes of students and teachers toward school were the dependent variable

	Variable	Variable heading
Student attitudes	Identification and connectedness to schools	School satisfaction (Bowen et al. 2000)
		Student school attachment (Crosnoe et al. 2004; Holas and Huston 2012; Kirkpatrick Johnson et al. 2001)
		Sense of belonging (Kahne, Sporte, De la Torre and Eaton)
		Achievement motivation (Koth et al. 2008)
		School connectedness (McNeely et al. 2002; Van der Vegt et al. 2005)
		Student engagement (Silins and Mulford 2004)
		Students sense of community in the school (Vieno et al. 2005)
		Classroom climate (De Winter 2003)
	Relationship with peers	Student engagement (Silins and Mulford 2004)
		Students sense of community in the school (Vieno et al. 2005)
		Relationships with peers (Van der Vegt et al. 2005)
	Relationship with teachers	Teacher support (Bowen et al. 2000)
		Student–teacher bonding (Crosnoe et al. 2004)
		Student school attachment (Holas and Huston 2012)
		Academic personalism, classroom personalism, student–teacher trust (Kahne et al. 2008)
		School connectedness (McNeely et al. 2002)
		Student engagement (Silins and Mulford 2004)
		Students' sense of community in the school (Vieno et al. 2005)
		Relationships with teachers (Van der Vegt et al. 2005)
Teacher attitudes	Identification and connectedness to schools	Teachers' collective responsibility (Lee and Loeb 2000)
		Communal school organization (Payne 2012)
		Organizational commitment (Rosenblatt 2001)
	Relationship with teachers	Teacher-teacher trust (Kahne et al. 2008)
		Communal school organization (Payne 2012)

Table 3.7 Results of vote counts examining the number of negative, nonsignificant, curvilinear, and positive effects of school size on students' and teachers' attitudes to school

	Studies	Samples	Direction of effect			
			−	*ns*	∩	+
School size measured as a continuous variable	9	9	12	4	0	0
School size measured as a quadratic function	1	1	0	1	1	0
School size measured as discrete variable (categories)	4	4	5	0	1	0
Total	14	14	17	5	2	0

− = negatively related with school size
ns = no significant relation with school size
∩ = optimal school size found
+ = positively related with school size

(students' participation or engagement). The latter refers to behaviors that represent participation, such as trying to their best in class, doing homework, and participate in extra-curricular activities. The authors further state that "theoretically, engagement and attachment are related to each other and to achievement. A student who feels more embedded in his or her school is more likely to exert effort, while one who participates in school and classroom activities is more likely to develop positive feelings about his or her school" (p. 320). Also, in previous research a positive relationship was found between identification and connection with aspects of schooling on the one hand and higher achievement and lower levels of problem behaviors on the other (e.g., Newmann et al. 1992; Bryk and Thum 1989; Gutman and Midgley 2000).

In this section, where the attitudes of students and teachers toward school are the outcome variable, we limit ourselves to attitudes to identification of and connection with school. Participation is addressed both in the section on involvement and in the section on other student outcomes.

Table 3.7 gives an overview of the number of studies, samples, and estimates included in the vote counting procedure for students' and teachers' attitudes to school. In total, 14 studies and 15 samples were included in the vote count. Two-third of the effects (derived from half of the 15 samples) between school size and attitudes to school appeared to be negative.

Two studies reported nonsignificant effects (Holas and Huston 2012; Kirkpatrick Johnson et al. 2001). Mixed effects were found in the studies by Crosnoe et al. (2004), Kahne et al. (2008), Van der Vegt et al. (2005). In these studies, both negative and nonsignificant effects were reported (see Tables A.8, A.9, A.10).

3.5.7 School Size Measured as a Continuous Variable

Eight studies reported linear negative effects of school size on attitudes to school. Five of these studies were conducted in the US, the other three in Australia, Israel, and Italy.

One the five US studies in which a negative effect was found is the study by McNeely et al. (2002). The authors used evidence from a sample taken from the National Longitudinal Study of Adolescent Health (about 75,000 adolescents from 127 schools, grades 7–12). Average level of school connectedness of pupils was the dependent variable. This variable measured the degree to which students felt close to people at this school, felt safe, and felt part of the school, were happy and experienced that the teachers treated them fairly. Multilevel analysis was employed. Variables included in the model were student background characteristics at individual and school level, teacher qualifications, structural school characteristics, discipline policies and student participation, and classroom management. The results showed that small school size is positively associated with school connectedness, but the strength of this relationship was meager, as an increase of 500 students in school size was associated with a very small decline in school connectedness.

The studies not conducted in the United States focused on respectively the impact of school size on teachers' organizational commitment in Israeli schools (Rosenblatt 2001), student engagement and participation in Australia (Silins and Mulford 2004) and students' sense of community in the Veneto region in Italy (Vieno et al. 2005). Negative effects of school size on students' attitudes and teachers' attitudes were reported in the studies, respectively bySilins and Mulford (2004) and Rosenblatt (2001). Vieno et al. (2005) found a positive effect, although this effect was not significant. In this latter study, conducted in the Italian context, students' sense of community was measured by a six-item scale (example items were "our school is a nice place to be, our students accept me as I am and when I need extra help I can get it from my teacher"). School size appeared to be non-significant in this study, as well as many of the other structural characteristics (e.g., facilities, extracurricular activities, and whether the school is public or private). SES was significant at the school level but not at the individual level. An intermediate variable positively associated with sense of community was democratic school climate, a variable better malleable to change than school size and other structural variables.

Silins and Mulford (2004) employed path modeling to examine the association between school size and SES on both students' perceptions of teachers' work in the class and students' outcomes (such as attendance, participation in, and engagement with school). Engagement with school was operationalized as students' perceptions with regard to the way teachers and peers relate to them, the usefulness of their schoolwork in later life, and the extent of identification with their school. School size had an indirect and negative effect on engagement through participation (i.e., absences, participation in extracurricular activities,

preparedness to do extra school work, involvement in classroom decisions, etc., ES = −0.16). Students in large schools participated less and this was associated with less engagement.

In the study conducted in the Netherlands finally, mixed effects were found. Van der Vegt et al. (2005) reported a nonsignificant effect of school size on students' connectedness with school and significant negative effects of school size on both relationships with peers and relationships with teachers.

3.5.8 Curvilinear Relationships

Like, McNeely et al., Crosnoe et al. (2004) also used data from the National Longitudinal Study of Adolescent Health. The sample included 15,000 students from 84 schools. The mean school size was 1,381 (with a standard deviation of 838). Interpersonal climate was the dependent variable. It was measured with three variables: (1) student school attachment (the extent to which adolescents felt close to people at their school and felt a part of their schools), (2) student–teacher bonding (the extent to which adolescents believed that teachers treated students fairly and, felt that teachers cared about them), and (3) student extra-curricular participation. Multilevel modeling was applied to estimate the effects of school size. The intra-class correlation (amount of variation between schools) appeared to be smaller for school attachment and teacher bonding (3 and 5 %, respectively) than for extra-curricular participation (14 %). For school attachment and teacher bonding a curvilinear effect was found with the lowest levels of attachment and teacher binding occurring at a size of 1,900 or 1,700 students, respectively. For extracurricular participation, a negative linear effect was found. The authors conclude that, based on the results of their study, an optimal school size for school connectedness would be less than 300 students, considerably lower than the optimal size for academic achievement found in other studies.

3.5.9 School Size Measured in Categories

In two of the tree studies in which school size was measured in categories (Bowen et al. 2000; Lee and Loeb 2000) small schools were favored above larger schools. In the study by Bowen et al., the focus was on student attitudes. School satisfaction and teacher support were the dependent variables. In the study by Lee and Loeb the impact of school size on teachers' collective responsibility was investigated by means of teacher attitudes, i.e., the extent of a shared commitment among the faculty to improve the school so that all students learn.

Bowen et al. conducted their study in middle schools in the US and used five size categories (the smallest 0–399 pupils, the largest 1,000–1,399 pupils). They found negative effects of school size on school satisfaction and teacher support and

Table 3.8 Overview of outcome variables and variable heading used in studies in which participation of students, teachers or parents was the dependent variable

	Variable	Variable heading
Participation of students	Extracurricular participation	Extracurricular participation (Coladarci and Cobb 1996; Crosnoe et al. 2004; Feldman and Matjasko 2006; Lay 2007; MacNeal 2008)
	Broader school participation	School involvement including school activity participation (Holas and Huston 2012)
		Participation in school activities (Silins and Mulford 2004)
Participation of teachers	Involvement in school decision making	Teacher influence (Kahne et al. 2008)
Participation of parents		Parent(s) act as a volunteer at the school (Dee et al. 2007)
		Average of total number of California Parent Teacher Association members for each affiliated school (Gardner et al. 2000)

concluded that "schools with enrolments of 800 or more might be too large to ensure a satisfactory educational environment." Lee and Loeb (2000) employed their study in 264 schools in Chicago. They found that compared to small schools (0–400 pupils) "teachers' views about the prevalence of collective responsibility appeared to be more negative in medium-sized schools (400–750 pupils) and even more in large schools (more than 750 pupils)".

De Winter (2003) also used three size categories in his study (less than 500, 500–1,000, more than 1,000 pupils), which was conducted in Dutch secondary education. He concluded that an optimal size, as far as school climate for pupils is concerned is that a school is neither too big nor too small.

3.5.10 Participation

Participation of students, teachers, or parents was the dependent variable in 10 studies (see Tables 3.9, A.11, A.12). With the exception of the study by Holas and Huston, in which primary and middle schools were sampled both, all other studies were concerned with secondary education. Nine studies were conducted in the United States and one in Australia (Silins and Mulford 2004).

Seven of the ten studies provided evidence on participation of students, one about teachers and two about participation of parents (see Table 3.8). In five studies, students' participation was restricted to participation in extracurricular activities; in the remaining two studies a broader operationalization of participation was taken. In the study by Holas and Huston, school involvement included four aspects (school attachment, teacher support, negative affect toward school and school activity participation). Higher scores represented higher involvement. Silins

Table 3.9 Results of vote counts examining the number of negative, nonsignificant, curvilinear, and positive effects of school size on participation

	Studies	Samples	Direction of effect			
			−	*ns*	∩	+
School size measured as a continuous variable	7	8	8	0	0	0
School size measured as discrete variable (categories)	4	5	2	2	1	0
Total	10	13	10	2	1	0

− = negatively related with school size
ns = no significant relation with school size
∩ = optimal school size found
+ = positively related with school size

and Mulford used a broad concept of students' participation, including absences, participation in extracurricular activities, preparedness to do extra work, involvement in classroom/school decisions and setting own learning goals, and voicing opinion in class.

The study by Kahne et al. (2008) examined the impact of 4 years of small school reform in Chicago. A variety of teacher and student measures was included in the study, including teachers' involvement in school decision making (see also the section on other dependent variables).

The impact of school size on participation of parents was examined in two studies. Dee et al. (2007) included four-dependent variables about parental involvement in their study, each variable measured through one single item. The item addressing the most intense involvement with school (i.e., volunteering at school) was chosen to be included in this review.

The results of the vote count for school size on participation are presented in Table 3.9.

In almost all samples a negative and significant association between size and participation was found despite different conceptualizations, outcome measurements, and types of respondents (see also Tables A.13, A.14). Although the number of studies is limited such a pattern of results supports the claim that smaller schools are associated with greater engagement. This was also found in other review studies (see Leithwood and Jantzi 2009).

A dissenting opinion came from the study by Lay (2007), titled "Smaller isn't always better." In this study data from the 1999, National Household Survey was used to examine the effects of school size on participation in school activities. School size was measured in three ways: based on parental answers about the enrollment of their child's school (responses were classified in categories) as well as based on data taken from matching zip codes for each pupil respondent with the high school within its borders (data were both used to measure school size as continuous variable as well as classified in categories). Other variables in the model were race, parent income, and plan to attend college. Depending on the measurement of school size used the effects on school activities differed. In the model where school size categories were based on parental responses (with

categories <300, 300–599, 600–999 and over 1,000) the association between school size and participation was nonsignificant. When school size was measured by a continuous variable (based on matching zip codes with each pupil response) the effect was significant and negative. Finally, when categories based on the continuous measure were used (with categories <300, 301–600, 601–900, 901–1,200, 1,501–1,800, and over 1,800) a curvilinear relationship was found, in which participation in schools with 1501-1800 students was significantly less likely. According to the author, concerns over the measurement of school size as well the limited number of student, school, and community variables included in the model may account for the few significant effects found.

Teacher influence was just one of the 10 teacher measures included in the multilevel models of a study on the implementation and impact of Chicago High School Redesign Initiative (CHSRI) (Kahne et al. 2008). In this initiative, large traditional neighborhood high schools (non-CHSRI schools) were converted into small autonomous ones. Data were collected for four successive waves of 11th graders starting in the 2002–2003 school year when three CHSRI conversion schools had 11th graders to the 2005–2006 school year when 11 CHSRI schools had 11th graders. Based on the theory of change ten teacher outcome variables (e.g., collective responsibility, quality professional development, teacher-teacher trust) were included in the study as well as ten student outcome variables (e.g., quality of English instruction, academic press, sense of belonging), and four outcomes (absences, drop-out rate, graduation rate, and achievement test scores). A great number of student and school level background variables were controlled for. Three level hierarchical linear modeling was used to estimate the significance and effects of the CHSRI schools compared to around the rest of the Chicago Public Schools (the non-CHSRI schools). The main conclusion is that "given the newness of the reform and the small size of the samples, it is clearly too soon to make broad claims about the efficacy of small school conversions in Chicago. … We see indications that small school conversions as promised provide a more personalized and supportive school context for students … We saw evidence that smaller schools enable the creation of contexts for teachers (e.g., ones characterized by greater trust, commitment. and sense of influence) but that these contexts do not appear to be fostering more systematic efforts at instructional improvement, different instructional practices, and improved performance on standardized tests" (p. 299).

3.5.11 School Safety

Evidence about the relationship between school size and school safety was derived from 24 studies (25 samples) (see Tables 3.11, A.15, A.16). Two studies were conducted in primary education (Bonnet et al. 2009; Bowes et al. 2009), one study used samples both from primary and secondary school students (O'Moore et al. 1997) and in three studies elementary and secondary school students were sampled

Table 3.10 Overview of outcome variables and variable heading used in studies in which safety was the dependent variable

Variable	Variable headings	Author(s)
Disciplinary school and class climate	School climate, respectful classroom behavior	Inspectorate of Education (2003), Kahne et al. (2008), Koth et al. (2008)
	Feelings of safety	Mooij et al. (2011)
	Students' behaviors (fights, use of alcohol, students' physical and verbal abuse of teachers etc.)	Bowen et al. (2000), Haller (1992)
	Misbehavior (disorder and bullying)	Chen (2008)
	School misbehavior	Stewart (2003)
Bullying	Bullying others and being bullies	Bowes et al. (2009), Klein and Cornell (2010), O'Moore et al. (1997), Van der Vegt et al. (2005), Wei et al. (2010), Winter (2003)
Problem behavior	Norm violating behaviors, alcohol, and marijuana	Chen and Vazsonyi (2013), Van der Vegt et al. (2005)
	Substance abuse while at school	Eccles et al. (1991)
	Suspensions	Heck (1993)
Violence	Sexual harassment	Attar-Schwartz (2009)
	Violence	Eccles et al. (1991), Leung and Ferris (2008), Van der Vegt et al. (2005), Watt (2003)
	Victimization (personal, property, physical, verbal)	Bonnet et al. (2009), Gottfredson and DiPietro (2011), Khoury-Kassabri et al. (2004), Klein and Cornell (2010)
	Crime (incidents)	Chen (2008), Chen and Weikart (2008)

together. The remaining 18 studies were conducted in secondary education. Thirteen studies were performed in the United States, five studies in The Netherlands (Bonnet et al. 2009; Inspectorate of Education 2003; Mooij et al. 2011; Van der Vegt et al. 2005; De Winter 2003), two in Israel (Attar-Schwartz 2009; Khoury-KassabrI et al. 2004), one in Ireland (O'Moore et al. 1997), one in the United Kingdom (Bowes et al. 2009), one in Canada (Leung and Ferris 2008), and one in Taiwan (Wei et al. 2010).

The outcome variables addressed in the 24 studies referred to various forms of student safety behavior, including (combinations of) disciplinary behavior, bullying, norm violating behavior, and different types of violence (see Table 3.10).

The summary of directions of effect for school size and safety is presented in Table 3.11 (for detailed information we refer to the Appendix, Tables A.17 and A.18). The results indicate that the number of negative and nonsignificant effects do not differ from each other.

Table 3.11 Results of vote counts examining the number of negative, nonsignificant, curvilinear, and positive effects of school size on safety

	Studies	Samples	Direction of effect			
			−	*ns*	∩	+
School size measured as a continuous variable	17	17	19	17	0	5
School size measured as discrete variable (categories)	8	9	3	5	2	3
Total	24	25	21	22	2	9

− = negatively related with school size
ns = no significant relation with school size
∩ = optimal school size found
+ = positively related with school size

3.5.12 Positive Relationships/Mixed Effects

Positive effects of school size on feelings of safety were reported in five studies. With the exception of the study by O'Moore et al. (1997) in which a sample from primary and secondary schools was taken, all studies were conducted in secondary schools. The findings suggest that pupils felt more safely in large schools (Mooij et al. 2011); that less bullying and fighting takes place in larger schools (Klein and Cornell 2010; O'Moore 1997; De Winter 2003), and that in larger schools pupils were more satisfied with the safety policy and regulations(Van der Vegt et al. 2005). In contrast to the findings of De Winter, Van der Vegt et al. reported a negative effect of size on bullying and fighting. The three Dutch studies (Mooij et al. 2011; Van der Vegt et al. 2005; Winter 2003), and the US study (Klein and Cornell 2010) will be discussed below, the study by O'Moore in the section on curvilinear relationships.

Mooij et al. (2011) used data from almost 80,000 pupils, 6,000 teachers, and other staff and 600 managers from secondary school in the Netherlands to test a two level model of social cohesion influences on a pupil's feelings of school safety. Personal background, level of attainment in education, school measures against violence (pro-social discipline) were positively associated with feelings of safety at school. Negative directions of effect were associated with not feeling at home in the Netherlands, peers taking drugs and weapons into school, by pupil's experiencing social violence, severe physical violence, and sexual violence as well as by staff experiencing severe physical violence. Curriculum differentiation based on learning differences (the streaming process of pupils into secondary schools) also had a negative effect on feelings of safety. The effect of school size was positive: pupils felt more safely at larger schools. However, when interaction effects were added to the model (i.e., the interaction of school size with pupil social violence), the main effect for school size on pupil's feelings of safety became insignificant. The authors conclude that "given the present results national policy should try to increase the safety of pupils and staff in school by enhancing pro-social rules of conduct and the shard control of these rules, taking school measures against truancy and redefining curriculum differentiation procedures" (p. 385/386).

Van der Vegt et al. (2005) investigated the effect of school size on feelings of safety, the availability of a safety policy, and the occurrence of bullying and fighting and vandalism, drugs and theft. About 5,000 secondary school pupils participated in the survey. Regression analysis was applied. The results found were both negative (more bullying and fighting, vandalism, drugs and theft at larger schools) and positive (pupils in large schools more satisfied with the safety policy and safety measures). School size had no effect on the perceptions (feelings) of safety.

De Winter (2003) found opposite effects, in this study being bullied, bullying and fighting occurred significantly more at smaller secondary schools, also after correction for level of attainment (school type, i.e., different streams of secondary education) or urbanicity. According to the author, an explanation might be that, as students at smaller schools do have more intense relationships with their peers, then more frequent bullying and fighting obviously might also be part of these contacts.

The study by Klein and Cornell (2010) is the only one of the 13 US studies that found positive effects. In this study, the data were collected in three different ways, by means of (1) student and teacher perceptions of victimization, (2) student self-reported number of experiences with victimization, and (3) rates of victimization based on school discipline records. Three types of victimization were the dependent variable (i.e., bullying, threats, and physical attacks). Other variables included in the model were poverty, proportion nonwhite students, diversity, and urbanicity. Regression analysis was applied. The results were mixed. When teacher and student perceptions of victimization were the dependent variable, the results indicated a negative effect (with significant higher levels of violence perceived in larger schools). However, nonsignificant effects were found when student self-reports of being a victim of violence were used. And if discipline violence rates were the measure, the results indicated a positive association. These contradictory findings suggest the need for a closer examination of the measures of victimization used: "If large schools truly have a higher rate of student victimization, it will be necessary for these schools to adopt stronger safety policies and prevention issues, but if the problem is one of perception only, then school authorities should focus on educational efforts to reassure students and help them to feel safe" (p. 943).

3.5.13 Negative Relationships

An inverse relation between school size and safety was reported in 11 studies (Attar-Schwartz 2009; Bowen et al. 2000; Chen 2008; Chen and Vazsonyi 2013; Eccles et al. 1991; Leung and Ferris 2008; Stewart 2003; see also Bowes et al. 2009; Gottfredson and DiPietro 2011; Haller 1992; Van der Vegt et al. 2005).The effect might be small (with an increase of e.g., 500 pupils in a school increasing the risk for being a victim of bullying after controlling for neighborhood and family background variables and children's internalizing and externalizing behaviors, see

Bowes et al. 2009), or discontinue, i.e., school size only matters for schools of a certain size category (see Leung and Ferris 2008).To explain evidence on the association between school size and safety in some studies it was argued that other school organization conditions than size might be more likely to influence safety (see Stewart 2003).

Leung and Ferris (2008) examined the effect of school size on self-reported teenage incidence of violence of 17-year-old low SES French speaking males in Montreal, Canada, controlling for social and demographic characteristics. School size was measured both continuously and classified into four size categories (1,000 or less, 1,000–1,499, 1,500–1,999, 2,000 or more). Control variables included in the binary logistic model were drop-out status, average family income at school level, family structure, delinquent friends, and parent's education. Depending on the measure of school size used, the results of the logistic regression analysis differed. School size measured continuously was significantly (negatively) associated with teenage violence. The authors also calculated marginal effects. For school size in the continuous model this implied that "an increase in school enrolment of one thousand would lead to about a 10 % increase in the probability of teenage violence" (p. 328). When school size was measured discretely (broken down into four size categories) only for very large schools a negative effect was indicated. "It's marginal effect suggests that teenagers who attended a school with more than 2,000 students were about 22 % more likely to engage in violent behavior than those who attended schools with less than 1,000 students" (p. 328). No significant effects were found for small -and large medium-sized schools.

School delinquency/misbehavior was the dependent variable in the study conducted by Stewart (2003). In this study, data were used from the second wave of the National Education Longitudinal Study (NELS). More than 10,000 10th grade students within 528 schools participated in the study. School misbehavior was measured by means of a scale asking pupils how often during the first half of the current school year they got in trouble for not following school rules, were put on an in-school suspension, suspended, or put on probation from school and got into a physical fight at school. Multilevel modeling was applied to examine the effects six of school level and 14 pupil level covariates on school misbehavior. Two school level variables in the model were significant: school size and school location. Larger schools in urban areas had significantly higher levels of school misbehavior. At individual level 10 of the 14 covariates were found significant, including three of the four school social bond variables distinguished in the study. Higher levels of school attachment, school commitment, and beliefs in school rules were positively associated with lower levels of misbehavior. School involvement, the 4th social bond variable, was (positive but) not significantly related to misbehavior. A further interesting result of this study is that the other school covariates (school composition, school poverty, school social problems, and social cohesion) were not significantly associated with school misbehavior.

3.5.14 Curvilinear Relationships

The only study that reported curvilinear relationships was the study by O'Moore et al. (1997). This study was conducted in Ireland in both a sample of primary and secondary schools. Three categories of size were distinguished (less than 200 students, 200–499 pupils, and 500 pupils or more). The results were mixed. In primary schools no significant differences were found between school size categories and the incidence of being bullied, while in secondary schools the chance of being bullied was least common in large schools. With regard to bullying others, both in primary and secondary education the highest proportion of pupils who bullied others were found in medium-sized schools.

3.5.15 Moderator Analyses

For the studies and samples in which school size was measured as a continuous variable moderator analyses were conducted to examine study and sample characteristics that may account for the differences of directions of effect found (see Table 3.12). The statistical technique employed and if a study was conducted in the United States are the most prominent outcomes. More negative effects are found in studies applied in the United States, as well as in studies that did not apply multilevel modeling. More significant effects (both negative and positive) were found if urbanicity was controlled for.

3.5.16 Student Absence and Dropout

Twelve studies (15 samples) reported on evidence about attendance, truancy, or absenteeism. The effect of school size on dropout was examined in four studies (5 samples). Almost all studies (and samples) were conducted in secondary schools, with one study reporting evidence from primary schools (Durán-Narucki 2008) and two studies employed in samples of both primary and secondary students (Eccles et al. 1991; Heck 1993). With the exception of the study by Bos et al. (1990), conducted in the Netherlands and the study by Foreman-Peck and Foreman-Peck (2006) conducted in Wales (United Kingdom), all studies relate to the context of the United States. Two studies (Gardner et al. 2000; Kahne et al. 2008) investigated the effect of size on both absenteeism and dropout.

The predominant outcome variables included in the studies were attendance, absenteeism, and drop-out rate (see Tables 3.13, 3.14, A.19, A.20, A.21, A.22). Perceptions with regard to truancy and absenteeism were measured in just two studies.

Table 3.12 Results of moderator analyses examining the number and percentage of negative, nonsignificant, and positive effects of school size on safety

Moderator	Negative effects N	Nonsignificant effects N	Positive effects N	Negative effects %	Nonsignificant effects %	Positive effects %
Level of schooling						
Primary school	1	2	0	33	66	0
Primary and secondary school	3	0	0	100	0	0
Secondary school	15	14	5	44	41	15
Country						
Canada	1	0	0	100	0	0
Israel	1	4	0	20	80	0
Netherlands	2	1	2	40	20	40
Taiwan	0	2	0	0	100	0
UK	1	2	0	33	67	0
USA	14	8	3	54	33	13
Covariates included						
Included covariate for SES	9	12	4	36	48	16
Included covariate for composite SES	14	14	3	45	45	10
Included covariate for urbanicity	8	3	4	53	20	27
Statistical technique used						
Technique multilevel	3	9	1	23	69	8
Technique not multilevel	16	8	4	57	29	14
Total	19	17	5	46	42	12

Before calculating the vote counts, the results of some studies were rescored, so that in all cases a positive effect denotes a situation of high attendance and less absenteeism, truancy or drop-out.

Table 3.15 shows the summary of the vote counts for studies in which attendance or truancy were the dependent variable. One study (Durán-Narucki (2008) reported a positive relationship between school size and attendance rate. Four studies reported negative effects (less attendance in larger schools) (Eccles et al. 1991; Foreman-Peck and Foreman-Peck 2006; Haller 1992; Jones et al. 2008). Mixed effects were reported in three studies (Kahne et al. 2008; Kuziemko 2006; Lee et al. 2011) and nonsignificant relationships in three studies as well (Bos et al. 1990; Chen and Weikart 2008; Heck 1993). One study (Gardner et al. 2000) reported evidence favoring small schools (see also Tables A.23, A.24).

With regard to drop-out, three of the five studies reported significant differences between size categories. In the fourth study (Kahne et al. 2008), in which a linear effect of size was investigated, no statistically significant relations were found (see also Table 3.16, A.25, A.26).

Table 3.13 Overview of outcome variables and variable heading used in studies in which attendance/absenteeism and truancy are the dependent variable

Variable	Variable headings	Author(s)
Truancy	Percentage of pupils absent	Bos et al. (1990)
	Perceptions with regard to truancy	Haller (1992)
Attendance	Attendance rate	Chen and Weikart (2008), Durán-Narucki (2008), Foreman-Peck and Foreman-Peck (2006), Heck (1993), Jones et al. (2008), Kuziemko (2006), Lee et al. (2011)
Absenteeism	Absenteeism rate	Gardner et al. (2000), Kahne et al. (2008)
	Perceptions with regard to absenteeism	Eccles et al. (1991)

Table 3.14 Overview of outcome variables and variable heading used in studies in which dropout is the dependent variable

Variable	Variable headings	Author(s)
Drop-out	Drop-out rate	Gardner et al. (2000), Kahne et al. (2008), Lee and Burkam (2003), Rumberger and Palardy (2005)

Table 3.15 Results of vote counts examining the number of negative, nonsignificant, curvilinear, and positive effects of school size on attendance/absenteeism and truancy

	Studies	Samples	Direction of effect			
			−	*ns*	∩	+
School size measured as a continuous variable	11	15	9	11	0	2
School size measured as discrete variable (categories)	1	1	1	0	0	0
Total	12	16	10	11	0	2

− = negatively related with school size
ns = no significant relation with school size
∩ = optimal school size found
+ = positively related with school size

Table 3.16 Results of vote counts examining the number of negative, nonsignificant, curvilinear, and positive effects of school size on drop-out

	Studies	Samples	Direction of effect			
			−	*ns*	∩	+
School size measured as a continuous variable	1	2	0	2	0	0
School size measured as discrete variable (categories)	3	3	1	0	2	0
Total	4	5	1	2	2	0

− = negatively related with school size
ns = no significant relation with school size
∩ = optimal school size found
+ = positively related with school size

3.5.17 Positive Relationships/Mixed Effects

Durán-Narucki (2008) investigated the relationship between the quality of school building facilities and poor English Language Arts and math achievement (i.e., percentage of students that scored on the lowest level) in 95 elementary schools in New York City. Attendance, measured as the average percentage of days attended school in a school year, was included as a potential mediator variable in the study. Covariates in the model were concentrated ethnicity, SES, teacher quality, and school size. The findings of the regression analysis indicated that school size was significantly and positively related with daily attendance, i.e., the study found significantly higher attendance in larger schools. The effects of school size on the percentage of students having poorer performance in English and math achievement were negative, but did not reach statistical significance. School attendance mediated the relation between school building condition and achievement, fully for poor performance in English Language Arts and partially for math. The author did not provide an explanation for the effect of size found in the study.

Lee et al. (2011) investigated the effectiveness of the Ohio High School Transformation Initiative (OHSTI) on attendance, graduation, dropout rates, and performance index scores. This school improvement initiative focused on transforming large high schools to small learning communities. In the Initiative a large school is defined as above 800 students, a small learning community as 100 students per grade level or 400 students in total. Between 30 and 35 schools participating in the study were small schools, approximately 200 schools were defined as large but being similar to the OHSTI schools. Mann-Whitney tests were performed to analyze attendance rates between small and large schools over 5 school years. In the first four years of the Initiative no significant differences were found between small and large schools, in the most recent school year (2007–2008) the attendance rate was significantly lower in small schools. Regarding drop-out rates (these were compared at grade level instead of school level and therefore not included in the review), the findings of the study indicated no consistent pattern. Although the study "observed some progress in small schools "the authors stated that "small schools programs alone are not the answer to improve education" (p. 25). Creating sense of community, extending the school day or year for students who need it and attracting and retaining effective teachers might be key factors as well.

3.5.18 Negative Relationships

Four studies reported negative effects (less attendance in larger schools). In two of these studies student and teacher ratings with regard to absenteeism were the outcome measure (Eccles et al. 1991; Haller 1992), while in the other two the effect of size on (attendance) rate was examined (Foreman-Peck and Foreman-Peck 2006; Jones et al. 2008).

Eccles et al. (1991) used data from the National Educational Longitudinal Study (NELS: 88). They found absenteeism, violence, and substance abuse significantly more often being reported as a major problem in larger schools by both teachers and students. Haller (1992) came to the same conclusion. In his study, perceptions of school level student indiscipline (truancy and vandalism/theft) was estimated from three sources (student, teacher, and self-reports) and regressed on school size and ruralness. The results show that ruralness and size together add significantly to the variance explained. Size appeared to be more important than ruralness. Interaction effects were also found: "the larger a rural school …, the greater its level of indiscipline" (p. 152). In the conclusion the authors hold a plea for other criteria than improving student behavior underlying decisions on consolidating schools (such as equity and efficiency). As far as student behavior is concerned, implementing relatively easy malleable school practices (such as identifying all pupils not attendant each morning) might be even effective as well.

3.5.19 Nonsignificant Relationships

Chen and Weikart (2008) investigated the relationship between school size, school disorder, student attendance, and achievement. The model builds upon the School Disorder Model (Welsh et al. 2000) and was extended for this study with student achievement. 212 middle schools in New York City participated in the study. Percentage free lunch and percentage white students were the control variables. Structural Equation Modeling was applied. Higher school disorder ($\beta = -0.10$), a lower attendance rate ($\beta = -0.08$), and lower achievement ($\beta = -0.02$) were found in larger schools but the effects were not statistically significant. The hypothesis that "school size has an indirect effect on academic achievement mediated by school disorder and student attendance rate" could also not be confirmed (p. 15). However, the results indicated a strong positive relationship between attendance rate and achievement ($\beta = 0.54$). Like Eccles et al., Chen and Weikart also suggest to focus on measures to improve school climate, including attendance policies, instead of reducing school size.

3.5.20 School Size Measured as Categories

Three studies reported differences on attendance or dropout rate between various school size categories (Gardner et al. 2000; Lee and Burkam 2003; Rumberger and Palardy 2005). Gardner et al. compared small Californian public schools (between 200 and 600 pupils) and large schools (2,000 pupils or more). Student achievement (four measures), absenteeism, and dropout were the dependent variables. The results indicated a significant positive effect of school size on all student achievement measures. At the same negative effects were found for absenteeism

and dropout. So students at larger schools performed better, but were more absent and dropout in large schools was significantly higher. This was also the conclusion in the study by Rumberger (1995). In this study (see the section on student achievement for a more elaborated description) an "inverted U" relationship was found for achievement and drop-out with large high schools (1,200–1,800 pupils) having significant higher achievement gain but also higher drop-out rates.

Lee and Burkam (2003) study built on the study by Rumberger (1995). Lee and Burkam also used the longitudinal data from the National Educational Longitudinal Study (NELS: 88). The sample consisted of 3,840 students in 190 schools from the High School Effectiveness supplement of NELS: 88. Whether a student dropped out between 10th and 12th grade was the dependent variable. Four categories of school size were compared (<600, 601–1,500, 1,501–2,500, >2,500). Binary logistic multilevel modeling was applied. The results indicated that "compared to medium-sized schools (601–1,500 pupils), large and very large schools have higher drop-out rates. This was particularly true for large schools (nearing a 300 % increase in the odds of dropping out, $p < 0.001$). Small schools also had higher dropout rates than medium-sized schools (more than a 100 % increase in the odds, $p < 0.10$)" (p. 22). Interaction effects indicated that in public or catholic schools of small and medium size with positive student–teacher relations, the probability on drop-out is less. The final model explains 12 % of the between school variance of drop-out. Besides the school level factors included in this study (school demographics, schools' academic organization, and schools' social organization) other factors might be of influence as well.

3.5.21 Other Student Outcome Variables

Six studies reported on school size effects on other student outcomes, i.e., student attitudes towards self and learning, and engagement (see Tables 3.17, A.27, A.28). One of these studies collected data from primary schools and middle schools (Holas and Huston 2012), the remaining studies all included evidence from secondary schools. One study (Inspectorate of Education 2003) was conducted in the Netherlands, the other six studies in the United States.

The results were mixed (see Tables 3.18, A.29, A.30). Two studies (Coladarci and Cobb 1996; Holas and Huston 2012) reported nonsignificant relationships between school size and student outcomes, two other studies reported negative effects (Lay 2007; Weiss et al. 2010). For one study (Kirkpatrick Johnson et al. 2001), a nonsignificant effect was found at the primary level, while at the secondary level larger schools were associated with less student engagement. In the study by Lay (2007) the direction of effect found differed depending on how school size was measured. When school size categories were the independent variable (either based on parental responses or on the continuous measure) a curvilinear relationship was found (with students in schools with fewer than 300 students significantly more likely to volunteer in community services). However, when school size was

Table 3.17 Overview of variables and variable heading used in studies on other student outcome variables

Variable	Variable headings	Author(s)
Attitudes	Pupil attitudes towards self or learning	Self-esteem (Coladarci and Cobb 1996)
		Perceived efficacy and competence in English and math (Holas and Huston 2012)
Behavior	Engagement	Engagement in school (Kirkpatrick Johnson et al. 2001)
		Academic engagement (Lee and Smith 1995)
		Participation in community services (Lay 2007)
		School engagement (Weiss et al. 2010)

Table 3.18 Results of vote counts examining the number of negative, nonsignificant, curvilinear, and positive effects of school size on other student outcome variables

	Studies	Samples	Direction of effect			
			−	ns	∩	+
School size measured as a continuous variable	4	5	2	3	0	0
School size measured as discrete variable (categories)	3	3	1	1	2	0
Total	5	7	3	4	2	0

measured continuously, the relationship between size and participation was nonsignificant.

3.5.22 Attitudes

Two studies, one in US middle and one in US high schools investigated the relationship between school size and student attitudes. Coladarci and Cobb 1996 examined the indirect effect of school size on 12th grade academic achievement and self-esteem through (total time spent on) extracurricular participation. Using evidence from the National Educational Longitudinal Study of 1988 database, only students who attended either a small high school (less than 800 pupils) or a large high school (1,600 or more pupils) were considered in the study. Structural equation modeling was applied. Variables included in the model were prior self-esteem and prior achievement, SES, size, total extracurricular participation and total time spent on extracurricular participation. The authors did find a significant negative effect of school size on extracurricular participation ($\beta = -0.210$), with higher extracurricular participation among students attending smaller schools. The indirect effects of school size on achievement ($\beta = -0.005$) and self-esteem ($\beta = -0.015$) through extracurricular participation were negative, but not significant.

Holas and Huston (2012) applied path analysis to compare student achievement, school engagement and perceived efficacy and competence in English and math of students starting middle schools in 5th and 6 grades compared to students of the same

grade in elementary schools. School characteristics (observed classroom quality, teacher-related classroom quality, school percentage of minority and poor students, and school size) were included in the path model as intermediate variables.The authors did not find significant effects of school size on any of the outcome variables of students in 5th grade. In 6th grade, school size was negative and significantly related to school engagement. In 6th grade, the study failed to find significant associations between size and perceived self-competence or achievement.

3.5.23 Engagement

Three studies investigated the impact of school size on student engagement in schools (Kirkpatrick Johnson et al. 2001; Lee and Smith 1995; Weiss et al. 2010). In these studies engagement in school was operationalized in very different ways (see Table A.27). Lee and Smith (1995) used the concept academic engagement, a composite of eight items measuring student behavior related to work in class. Kirkpatrick Johnson et al. (2001) focused on engagement in school (operationalized as attendance, attention for school work and doing homework), while Weiss et al. (2010) used a very broad composite measure of engagement based on seven variables: teacher experience, delinquent behavior, academic friend, educational motivation, teachers' belief about ability, school preparedness, and parental involvement.

Lee and Smith (1995) investigated the effects of school size on achievement gain and academic engagement, using data from the National Educational Longitudinal Study 1988. Their analysis controlled for school restructuring practices, SES, minority status, initial ability, average school SES, minority concentration, sector, academic emphasis, and course-taking differentiation. The authors found both significantly higher and more socially equitable achievement gain and academic engagement in smaller schools. In the discussion of the article the authors wonder whether reducing school size really is the issue. "We would not draw that conclusion from our results. … Rather the findings indicate that the size of enrolments act as a facilitating or debilitating factor for other desirable practices. For example, collegiality among teachers, personalized relationships, and less differentiation of instruction by ability … are more common and easier to implement in small schools" (p. 261/262).

Weiss et al. (2010) also investigated the impact of size on achievement and engagement in US high schools. Using data from the Educational Longitudinal Study (ELS 2002) they found that "there are significant differences related to student engagement between schools of different size categories, while school size is not significantly related to mathematics achievement. Compared with students attending schools of the smallest size (the omitted category in the multilevel analysis), those in schools with 1,000–1,599 students or with more than 1,600 students have (significant) lower levels of engagement" (p. 170). Differences related to demographic characteristics were also examined in the study. Students previously held back were significantly less engaged, students from higher

Table 3.19 Overview of outcome variables and variable heading used in studies on school organisation and teaching and learning

Variable	Variable headings	Author(s)
Teaching and learning	Expectations and support	Expectations for postsecondary education (Kahne et al. 2008)
		Academic press (Kahne et al. 2008)
		Peer support for academic achievement (Kahne et al. 2008)
		School-wide future orientation (Kahne et al. 2008)
	Instruction	Pedagogical and didactical approach (Inspectorate of Education 2003)
		Quality student discussions in classroom (Kahne et al. 2008)
		Quality English instruction (Kahne et al. 2008)
		Quality math instruction (Kahne et al. 2008)
		Teachers' work (Silins and Mulford 2004)
School organization	Teacher attitudes	Teacher efficacy (Eccles et al. 1991)
		Teachers' collective responsibility (Kahne et al. 2008)
		Commitment to innovation (Kahne et al. 2008)
	Leadership	Principal instructional leadership (Kahne et al. 2008)
		Teacher Leadership (Silins and Mulford 2004)
	Curriculum	Program coherence (Kahne et al. 2008)
	Professional development	Quality professional development (Kahne et al. 2008)
		Reflective dialogue (Kahne et al. 2008)
	Organizational learning	Organizational learning (Silins and Mulford 2004)

educated parents, students with higher SES, students with Hispanic background and females have significantly higher engagement. African–American students were not significantly different in engagement than white students.

3.5.24 School Organization and Teaching and Learning

Three studies in the review included measures of the impact of school size on school organization and teaching and learning (see Table 3.19). These studies had different aims and scope.

Thirteen of the 17 effects reported are derived from the study by Kahne et al. (2008), three from the study of Silins and Mulford, and each one from the study by Eccles et al. (1991) and the study of the Dutch Inspectorate of Education (see Tables A.31, A.32). The results of the vote counts are mixed: most effect sizes appeared to be not significant, six effects reported were negative (favouring small

Table 3.20 Results of vote counts examining the number of negative, nonsignificant, curvilinear, and positive effects of school size on school organization and teaching and learning

	Studies	Samples	Direction of effect			
			−	*ns*	∩	+
School size measured as a continuous variable	3	3	6	11	0	0
School size measured as discrete variable (categories)	1	1	0	0	1	0
Total	4	4	6	11	1	0

− = negatively related with school size
ns = no significant relation with school size
∩ = optimal school size found
+ = positively related with school size

schools), and for one study a curvilinear relationship was found (see Tables 3.20, A.33, A.34).

3.5.25 Negative and Nonsignificant Relationships

The study by Kahne et al. (2008) focused on the implementation and impact of the first phase of the Chicago High School Redesign Initiative (CHSRI). A theoretical framework summarizing the theory of change underlies this study and portrays the mechanisms through which the characteristics of small school reform are thought to promote a supportive and personalized context for students as well as a desirable teacher context for reform, which in turn would impact on instruction and different types of student outcomes (absences, drop-out rate, graduation rate, and achievement test scores) (for a more elaborated description see also the section on participation. The results of the three level multilevel analysis yielded four significantly negative effects and nine nonsignificant effects. It was found that teachers in CHSRI schools had a better context for reform (significantly greater level of commitment to innovation and a higher sense of collective responsibility). CHSRI schools also provided a more supportive context for students (with significantly higher expectations for postsecondary education and school-wide future orientation, but no significant difference for peer support for academic achievement). However, after the first phase, the improved contexts for teacher and students in CHSRI schools did not have a statistically significant impact on facilitators for instructional improvement(principal leadership, professional development, program coherence) and improved instructional practices (quality of student discussions, quality of English and math instruction, academic press). So although some significant positive indications of the effects Chicago High School Redesign Initiative were visible, after 5 years it still "might be too soon to make broad claims about the efficacy of small school conversions in Chicago" (p. 299).

Silins and Mulford (2004) employed path modeling to examine the impact of school external (size and SES) and school internal variables on teacher leadership,

organizational learning, teachers' work and ultimately students' outcomes (i.e., participation in and engagement with school). The study was conducted in Australia. School size had a significant negative indirect effect on organizational learning through staff perceptions of the availability of resources. School size was not significantly associated with teacher leadership and teachers' work.

3.5.26 Curvilinear Relationship

The study of the Dutch Inspectorate of Education (2003) had the aim to investigate the associations between various aspects of the quality of Dutch secondary schools as assessed by the Inspectorate (such as achievement, pedagogical and didactical approach, pupil guidance, and quality care) and elements of school structure (size, school types, and locations). In this study, a curvilinear effect was found between school size and the quality of the pedagogical and didactical approach. The results indicated mid-size schools (500–1,000 pupils) having the lowest score on the quality of the pedagogical and didactical approach.

3.5.27 Costs

The review on costs was limited to studies that investigated variations in per pupil expenditure between schools of different sizes. Studies in which costs were measured at the above school level (at the district level for example as in Chakraborty et al. (2000)) were excluded.

Five studies investigated variations in economic outcomes at school level (see Tables A.35, A.36). Four studies were from the USA and one from the Netherlands. Two studies were conducted in primary education (Merkies 2000; Stiefel et al. 2000), one in secondary education (Bickel et al. 2001) and two studies related to both primary and secondary education (Bowles and Bosworth 2002; Lewis and Chakraborty 1996).

All studies reported a significant negative effect of school size on costs per pupil (Bickel et al. 2001; Bowles and Bosworth 2002; Lewis and Chakraborty 1996; Merkies, Stiefel et al. 2000) (see Tables 3.21, A.37). A similar pattern was reported in each study. Sharp decreases in per pupil expenditure occur as the school size increases from very low to average, whereas the increase from average onwards is associated with much more modest decreases in costs. All studies take into account the impact of student population characteristics (e.g., income and ethnicity) and educational output (e.g., achievement scores, dropout or graduation rates) when assessing the effect of school size on costs per student. The effect of school size remains intact when controlling for educational output. In the study by Stiefel et al. (2000), however, the effect of school size largely disappears when taking into account student population characteristics (especially limited English proficiency).

Table 3.21 Results of vote counts examining the number of negative, nonsignificant, curvilinear, and positive effects of school size on costs

	Studies	Samples	Direction of effect			
			−	*ns*	∩	+
School size measured as a continuous variable	4	4	4	0	0	0
School size measured as discrete variable (categories)	1	1	0	1	0	0
Total	5	5	4	1	0	0

− = negatively related with school size
ns = no significant relation with school size
∩ = optimal school size found
+ = positively related with school size

3.5.28 Negative Relationships

Bickel et al. (2001) examined the association between size, achievement, and costs (expenditure per pupil) in 1,001 Texas high schools. Besides the effect of size on costs for the total group of schools, the authors were also interested in the differential effects for the two types of high schools that could be distinguished in the sample: "conventional high schools," schools serving a narrow range of secondary school grades, and "single-unit schools," schools typically the only school in a small rural district spanning all elementary and secondary grades. Other variables included in the study were ethnic, linguistic, and socioeconomic background of pupils, organizational and curriculum characteristics, achievement, and student–teacher ratio. The results of the multiple regression analysis indicated that school size was negatively related to expenditure per pupil, in total and also for conventional and single-unit schools. But compared to conventionally grade-specialized high schools, single unit schools were associated with substantial lower expenditure per pupil. On average, the savings in single unit schools correspond to a reduction of over $1,000 per pupil. The savings decline as these schools become larger. Bickel et al. attribute the savings to a diminished need for coordination and control due the facts that single unit school in all cases were the only school in the district, and covered the full range of grades.

Bowles and Bosworth (2002) used data that contained rather detailed expenditure data to examine the effect of size on expenditure per student across a 4-year period (1994–1998). Data were collected from 80 primary, middle, and high schools in Wyoming. The authors applied different regression models. The results were consistent, finding a negative effect across all model specifications, suggesting that the expenditure per pupil decreases as school size increases. Across school types, "an increase of 10 % in school size decreases costs per student by approximately 2 %" (p. 299).

Lewis and Chakraborty (1996) investigated the effect of both school size and district size on cost per student using data from Utah (U.S.). Their analyses controlled for educational output (dropout and graduation rates) and several other

relevant factors (e.g., income, teacher salaries). An inverse relation between school size and costs per student was established. The analyses also indicated that the impact of school size on costs per student clearly outweighs the impact of district size.

The fourth study (Merkies 2000) relates to primary school in the Netherlands. Here an optimal size of around 450 pupils is reported. It was found that "from the average school (200 pupils) onwards the average costs remain virtually constant. For schools with more than twice the average number of pupils there are hardly any more economies of scale" (p. 206).

The last study included in the review (Stiefel et al. 2000) estimated the effect of size on the budget per student and on the 4-year budget per graduate (a combined output and cost measure), while controlling for type of school and student background. 121 New York City public high schools participated in the study. Three categories of school size are compared (0–600, 600–2,000, >2,000 pupils), each including different types representing the mission or the program of the school. The authors reported a significant negative effect of school size on both budget per student and 4-year budget per graduate. When taking into account school population characteristics (especially limited English proficiency) differences in budget per graduate turned out to be minimal: "small schools are cost effective but so are also large schools in New York City" (p. 36–37).

3.6 Conclusion

In this chapter, the results of a research synthesis of the effects on school size on various outcome variables are presented. The research synthesis sought answers on the following research questions:

(1) What is the impact of school size on various cognitive and noncognitive outcomes?
(2) What is the "state of the art" of the empirical research on economies of size?

To answer the first question the impact of school size of variety of student, teacher, parents', and school organizational outcome variables was investigated. A distinction was made between outcome variables, i.e., cognitive and noncognitive outcome variables, and school organization variables. To answer the second question, costs was included as a dependent variable in the review.

A meta-analysis of the vote-count type was carried out, which means that an overview is given from studies and samples that showed significant positive, significant negative, curvilinear or nonsignificant relationships between school size and various dependent variables. Eighty studies, 127 samples, and 277 estimates were included in the vote counting procedure. The results are presented in Table 3.22.

The overall pattern of the vote counting procedure show that, across all studies that examined the association between school size and any dependent variables,

Table 3.22 Directions of effect of school size on various dependent variables

Dependent variable	Studies	Samples	Direction of effect							
			− N	*ns* N	∩ N	+ N	− %	*ns* %	∩ %	+ %
Achievement	46	64	23	78	14	11	18	62	11	9
Students' and teachers' attitudes to school	14	14	17	5	2	0	71	21	8	0
Participation	10	13	10	2	1	0	77	15	8	0
Safety	24	25	21	22	2	9	39	41	4	17
Attendance/absenteeism and truancy	12	16	10	11	0	2	43	48	0	9
Drop-out	4	5	1	2	2	0	20	40	40	0
Other student outcome variables (attitudes towards self and learning, engagement)	5	7	3	4	2	0	33	44	22	0
School organization and teaching and learning	4	4	6	11	1	0	33	61	6	0
Costs	5	5	5	0	0	0	100	0	0	0
Total[a]	84	107	95	136	23	23	35	49	8	8

− = negatively related with school size
ns = no significant relation with school size
∩ = optimal school size found
+ = positively related with school size
[a] Because publications and samples may refer to more than one dependent variable, the total number of publications and samples is lower than the sum of samples and publications

almost half (49 %) of the effect estimates appeared to be nonsignificant, and one-third (34 %) negative. Positive effect relationships were found for less than 10 % of the estimates. Based on these overall results we cannot conclude that smaller schools are generally better for all types of outcomes.

However, when attitudes of students and teachers toward school or participation of students or parents in school (related) activities were the outcome variables, the results tend to indicate a negative association. The operationalization of attitudes in the studies referred to identification and connection with school (both students and teachers), relationships with peers or colleagues and relationships with teachers (students). Participation was operationalized as either participation in school related or extracurricular activities (students), act as a volunteer or being member of a parent association (parents) and involvement in decision making (teachers). For attitudes and participation, 70 % or more of the estimates was negative, none positive, and for studies and samples in which nonsignificant effects were reported the direction appeared to be negative as well.

The relationship between size and academic achievement was investigated in more than half of the number studies included in the review. The results show a mixed pattern with 62 % of the associations between size and achievement reported as statistically nonsignificant, 20 % as negative and 9 % positive.

The pattern for safety and attendance and truancy show results that are comparable to the overall results. For safety and attendance the number of negative and nonsignificant findings do not differ that much from each other. However, on the contrary to what was found for attitudes and participation, where nonpositive effects were reported, for safety one out of five estimates were positive (17 % of the estimates, derived from five studies). In the studies that found positive effects, specific measured of safety were addressed. In these studies safety referred to either more general feelings (pupils felt more safely in large schools, Mooij et al. 2011); bullying and fighting (bullying and fighting occurred less in larger schools, Klein and Cornell 2010; O'Moore 1997; Winter 2003), and more satisfaction with the safety policy and regulations (Van der Vegt et al. 2005). Other operationalizations used in the studies, for which no positive effects were found, referred to (combinations of) disciplinary behavior, bullying, norm violating behavior, and several types of violence.

The association between school size and school organization and teaching and learning was investigated in three studies. The majority of effects reported (13 out of 17) are derived from one study. As for achievement the results found are mixed, with more than half of the estimates being nonsignificant.

For academic achievement and safety moderator analyses were carried out for the studies and samples in which school size was measured as a continuous variable. For academic achievement the most striking outcomes of these analyses concerned the statistical technique employed and the inclusion of a covariate for student's prior achievement in the model. Negative effects were more found in studies that account for prior achievement as well as in studies that employed multilevel modeling. For safety more negative effects were also found in applied multilevel modeling. Next to this, the percentage of negative effects found is somewhat higher for studies conducted in the US context and more significant (both positive and negative) effects were found if urbanicity was controlled for.

The review of costs was limited to studies that investigated variations in per pupil expenditure between schools of different sizes. All five studies included in the review reported a negative effect of school size on costs per pupil. The pattern reported in each study was in the same direction: sharp decreases in per pupil expenditure occur as the school size increases from very low to average, whereas the increase from average onwards is associated with much more modest decreases in costs.

Annex

Student Achievement

See Tables A.1, A.2, A.3, A.4 and A.5.

Table A.1 Overview of studies of school size on student achievement

Overview of studies of school size on student achievement

Authors	Sample	Country	School type[a]	School size measure	Student achievement measure	Mean	SD	Database
Åberg-Bengtsson (2004)		Sweden	P	Small rural schools: Schools with an enrolment of less than 75 students and located in a rural district versus schools with an enrolment of 75 students and more	Reading literacy			IEA reading literacy (1990–1991)
Alspaugh (2004)		USA (Missouri)	P	School size (K-5 enrolment) <200 200–299 300–399 400–499 >=500	Composite (5 Stanford 9 NCE achievement measures: reading, math, language, science, social science)			
Archibald (2006)		USA (Nevada, Washoe county school district in Reno)	P	The number of students enrolled at the school	Student level post-test scores reading and mathematics	548	137	
Barnes et al. (2006)	KS1 KS2	England (deprived areas)	P	Total number of students at the school roll	KS1 English, math KS2 English, math, science			Data collected as part of the National Evaluation of Sure Start (NESS 2004)
Bickel et al. (2001)		USA (Texas)	S	Number of students: expressed in thousand students units Expressed in natural logarithms of single-student units	Texas assessment of academic skills 10th grade Reading Math Writing Composite achievement	877	850	Texas dataset representing 1,001 high schools

(continued)

Table A.1 (continued)

Overview of studies of school size on student achievement

Authors	Sample	Country	School type[a]	School size measure	Student achievement measure	Mean	SD	Database
Borland and Howsen (2003)		USA (Kentucky)	P	School size number of students within a school School size squared: square of number of students within a school	The mean total battery normal curve equivalent score for 3rd grade students within a school (combined subject scores for reading, language and mathematics)	490	204	
Bowles and Bosworth (2002)		USA (Wyoming)	PS	Average daily membership for school *i* for period *t*	Arithmetic average of the three scores on standardized test results for reading, writing and math for each public school in Wyoming			Data from 17 Wyoming school districts
Bradley and Taylor (1998)	11–16 Schools 1992 1996	UK	S	Pupils/100 Pupils/100 squared	School exam performance (proportion of 15–16 year-old pupils in each school obtaining five or more GCSEs at grades A to C) (four performance categories)	685 765		Secondary school performance tables and information obtained from the annual schools' census undertaken by the department for education and employment
	11–18 Schools 1992 1996					916 1010		
Caldas (1993)	P S	USA (Louisiana)	P S	The number of students enrolled in the school in October 1989	Average school score composite of norm-referenced test and criterion-referenced test (language, math, writing, science and social studies)	507 683	223 384	Data collected and aggregated by the Louisiana State Department of Education

(continued)

Table A.1 (continued)

Overview of studies of school size on student achievement

Authors	Sample	Country	School type[a]	School size measure	Student achievement measure	Mean	SD	Database
Carolan (2012)		USA (nationally representative)	S	High school size: Small (<600 students) Moderate (600–999 students) Moderately large (1,000–1,599 students) Large (>1,599 students)	12th grade math score: a student score on the ELS math assessment			Education longitudinal study (ELS) 2002
Chen and Weikart (2008)		USA (New York City)	S	Number of students enrolled at each school	School mean score on the grade 8 English language arts and the grade 8 mathematics of the New York State Examinations for the 2003–2004 school year	960	493	Data from New York City Department of Education (2002–2003 and 2003–2004 school year data for all middle schools)
Coladarci and Cobb (1996)		USA (nationally representative)	S	School size Compares smaller (<800) and larger (>=1600) schools Students from schools with other sizes eliminated from analyses	NELS: 88 senior year composite of student performance in reading and mathematics			National Education Longitudinal Study of 1988 (NELS: 88)
Deller and Rudnicki (1993)		USA (Maine)	P	Average daily attendance for the year 1985	Achievement of grade 8 students: three year (1986–1987 through to 1988–1989) cumulative aggregate test score as compiled by the Maine educational assessment program (covering reading, writing, math, science, social studies and humanities)			Databases Maine department of education

(continued)

Table A.1 (continued)

Overview of studies of school size on student achievement

Authors	Sample	Country	School type[a]	School size measure	Student achievement measure	Mean	SD	Database
Driscoll et al. (2003)	Primary Middle High school	USA (California)	P S S	School size	1999 California academic performance index (weighted average of Stanford 9 test scores) (aggregated at school level)	526 526 526	394 394 394	California Department of Education database
Durán-Narucki (2008)		USA (New York City, Manhattan borough)	P	The number of students enrolled at each school	Poor achievement: percentage of students at each school that scored at level 1 in the New York State and New York City tests on English language arts and math	712	328	Building Condition Survey (comprehensive study on the condition of New York City school buildings School report cards for the year 2000 New York City Board of Education
Eberts et al. (1990)		USA (Nationally representative)	P	<200 (small) 58 schools 400–599 (medium) 94 schools >800 (large) 19 schools Categories 200–399 (86 schools) and 600–799 (30 schools) purposefully omitted	Gains in mathematics achievement			Subset of the sustaining effects study conducted by the system development corporation for the former Office of Education (1987)
Fernandez (2011)		USA (Nevada, Clark county school district)	PS	The number of students enrolled	Iowa test of basic Skills growth in math and reading (2005–2006) school performance score	1082	637	Dataset clark county school district Study explores the relationship between the quality of a school improvement plan and school performance
Foreman-Peck and Foreman-Peck (2006)		UK (Wales)	S	Log (previous year pupil numbers) School size 1996 School size 2002	% of pupils in a school gaining 5 or more A-C GCSEs	871 936	331 328	Dataset provided by the school and teacher statistics division of the Welsh assembly government

(continued)

Table A.1 (continued)

Overview of studies of school size on student achievement

Authors	Sample	Country	School type[a]	School size measure	Student achievement measure	Mean	SD	Database
Fowler and Walberg (1991)		USA (New Jersey)	S	Total enrolment for the 1984–1985 school year for all grades	Minimum basic skills test –Average reading test score –Average Math test score –% passing Reading test score –% passing Math test score –% Passing both reading and math score High school proficiency test –Average reading test score –Average Math test score –Average writing test score –Average writing test mc score –Average writing test essay score –% passing Reading test score –% passing Math test score –% passing Writing test	1070	519	Data obtained from the New Jersey department of education
Gardner et al. (2000)		USA (Cailfornia)	S	School size Small schools (200–600 pupils) Large schools (>2,000 pupils)	Scholastic aptitude test (SAT) Verbal SAT Math SAT	424 2500		Data were obtained from the 1995/1996 California high school performance report
Heck (1993)		USA (Western state)	PS	Actual size of enrolment	School reading and mathematics scores on the Stanford achievement test (compiled over a 2-year period. 1989–1991)			State department of education's survey on restructuring the curriculum
Holas and Huston (2012)	Grade 5 Grade 6	USA (nationally representative)	PS	Total enrolment	Tested achievement (reading math) Teacher-related achievement Teacher-related achievement	490 690 300	210	NICHD study of early child care and youth development

(continued)

Table A.1 (continued)

Overview of studies of school size on student achievement

Authors	Sample	Country	School type[a]	School size measure	Student achievement measure	Mean	SD	Database
Kahne et al. (2008)		USA (Chicago)	S	School size	Scores of the 11 grades on the Prairie State Achievement Exam in math and English			Consortium on Chicago school research's biannual survey. Administrative records of CPS and test data
Kuziemko (2006)		USA (Indiana)	P	Abrupt change in school enrolment	Absolute change in average ISTEP math score change over 1 year, 2 years, 3 years Absolute change in average ISTEP language score change over 1 year, 2 years, 3 years	418	170	Indiana department of education: Indiana statewide test for educational progress Public school universe data form national center for educational statistics
Lamdin (1995)		USA (Baltimore, Maryland)	P	The number of students enrolled in grade kindergarten through five	The percentage of students in each school above the mean reading score on the California achievement test The percentage of students in each school above the mean math score on the California achievement test	469	172	Data from 1990 report by the Baltimore citizens planning and housing association
Lee et al. (2011)	03–04 04–05 05–06 06–07 07–08	USA (Ohio)	S	Small schools in Ohio versus traditional schools that are identified as similar to the small schools Large school at or above 800 students. small learning communities approximately 100 students per grade level or 400 students with the learning community	Graduation rate Performance index score			Data collected from more than 230 Ohio schools

(continued)

Table A.1 (continued)

Overview of studies of school size on student achievement

Authors	Sample	Country	School type[a]	School size measure	Student achievement measure	Mean	SD	Database
Lee and Loeb (2000)		USA (Chicago)	P	Number of students in the school Categories: <400 (RF) (5.2 % of students) 400–750 (48.8 %) >750 (46.0 %)	1997 mathematics score: score for 6th and 8th graders combined in a single scale. Each student's grade equivalent score on the math section of the Iowa test of basic skills administered to all Chicago elementary school students each year			Data provided by the consortium on Chicago school research
Lee and Smith (1995)		USA (nationally representative)	S	Total enrolment of as October 1989 (transformed to its natural logarithm and standardized)	IRT estimated gain between 8th and 10th grade test (NELS) Mathematics Reading History Science			National education longitudinal study of 1988 (NELS: 88) 1st and 2nd wave Mean school size: Traditional schools: 1095 Moderate schools: 633 Restructuring schools: 764
Lee and Smith (1997)		USA (nationally representative)	S	Total enrolment as of October 1989 Categories: <300; 301–600; 601–900; 901–1,200; 1,201–1,500; 1,501–1,800; 1,801–2,100; >2,100	Achievement gain math between 8th and 12th grade Achievement gain reading between 8th and 12th grade			National education longitudinal study of 1988 (NELS: 88) 1st three panels
Lubienski et al. (2008)	Grade 4	USA (nationwide)	P	School enrolment Categories: 1–299, 300–499, 500–699, and 700 or more (1–4 scale)	4th grade mathematics achievement NAEP 2003			National assessment of educational progress (NAEP) 2003
	Grade 8		S	Categories: 1–399, 400–599, 600–799, 800–999, and 1,000 or more (1–5 scale)	8th grade mathematics achievement NAEP 2003			

(continued)

Table A.1 (continued)

Overview of studies of school size on student achievement

Authors	Sample	Country	School type[a]	School size measure	Student achievement measure	Mean	SD	Database
Luyten (1994)	USA 1st and 2nd sample	USA	S	School size 5 categories: <240, 240–359, 360-499, 500-999, >1,000 (1–5 scale)	75 item math test SIMS			Second international mathematics study
	Sweden	Sweden			75 item math test SIMS			Second international mathematics study
	Netherlands math	Netherlands math			75 item math test SIMS			Second international mathematics study
	Netherlands science	Netherlands science			61 item test physics, chemistry, biology and earth science SISS			Second international science study
Ma and McIntyre (2005)		Canada (Central Alberta)	S	School size: expressed in hundred students units	Mathematics subtest of the Canadian achievement test (2nd edition)	613	363	Longitudinal study of mathematics participation (exploring differential effects of mathematics courses in mathematics achievement) Interaction effect with course-taking: students taking pure math courses, students taking applied math courses
Maerten-Rivera et al. (2010)		USA (large urban school district southeast USA)	P	School size	5th grade science achievement (high-stakes test)	798	331	Data collected for all 5th grade students in a large urban school district in the southeast United States during the 2006–2007 school year

(continued)

Table A.1 (continued)

Overview of studies of school size on student achievement

Authors	Sample	Country	School type[a]	School size measure	Student achievement measure	Mean	SD	Database
McMillen (2004)	Primary school	USA (North Carolina)	P	4 categories: <400; 400–549; 550–699; >700	End of grade 3 test Reading Math	506		Several databases maintained by the North Carolina department of public instruction. data on school size from state student membership database
	Middle school		S	4 categories: <400; 400–549; 550–699; >700	End of grade 6 test Reading Math	570		
	High school		S	4 categories: (<700; 700–1,199; 1,200–1,699; >1,700	High school comprehensive test grade 8 Reading Math	859		
Moe (2009)	Elementary	USA (California)	P	The log of school enrolment	The growth in academic performance index (API) scores between 1998–1999 and 2002–2003 (API based on the scores of all students across all grades i all subjects tested)			Data drawn from school districts in the state of Calaifornia
	Secondary		S	The log of school enrolment	The growth in academic performance index (API) scores between 1998–1999 and 2002–2003 (API based on the scores of all students across all grades i all subjects tested)			Data drawn from school districts in the state of California

(continued)

Table A.1 (continued)

Overview of studies of school size on student achievement

Authors	Sample	Country	School type[a]	School size measure	Student achievement measure	Mean	SD	Database
Ready and Lee (2006)		USA	P	Categories: <275 (small) 276–400 (medium-small) 601–800 (medium-size)-(RF) 601–800 (medium-large) >800 (large)	ECLS-K Literacy: both basic literacy skills as well as more advanced reading comprehension skills (assessment administered individually) ECLS-K math: conceptual and procedural knowledge and problem solving, equally divided between number sense and measurement			Early childhood longitudinal study kindergarten Cohort (ECLS-K, first four data waves)
Rumberger and Palardy (2005)		USA (nationwide)	S	Categories: 1–600 (small) (23 %) 601–1,200 (medium) (36 %) RF 1,201–1,800 (large) (28 %) >1,800 (extra large) (13 %)	Achievement growth: mean of math, reading, science and history test scores administered in the spring semesters of 1988, 1990 and 1992 when most students were enrolled in grades 8, 10 and 12			National education longitudinal survey (NELS: 88)
Sandy and Duncan (2010)	Urban Suburban	USA (nationally representative)	S	Small schools (<1,000) Large schools (>1,000)	Composite of arithmetic reasoning, math, word knowledge, paragraph comprehension, general science, numerical operations, coding speed, mechanical comprehension, electronics information, auto and shop information (armed services vocational aptitude battery)			National longitudinal survey of labour market experience for youth (1997)

(continued)

Table A.1 (continued)

Overview of studies of school size on student achievement

Authors	Sample	Country	School type[a]	School size measure	Student achievement measure	Mean	SD	Database
Sawkins (2002)	1993–1994	UK (Scotland)	S	Total number of pupils/100 (Total number of pupils/100) squared	Examination performance: % of pupils in S4 gaining 5 or more standard grade passes at levels 1 or 2 (credit level)	796	356	Examination results in Scottish schools/annual statistical returns made to the Scottish executive
	1998–1999					806	356	
Schneider et al. (2006)		USA (nationally representative)	S	Total school enrolment Categories: 1–399 400–799 800–1199 (reference category) 1200–1999 2000 or more	12th grade mathematics achievement			Educational longitudinal study of 2002 (ELS: 2002)
Stewart (2008)		USA (nationally representative)	S	Total student enrolment of the school	Grade point average (based on current grades in math, English, history and science	1540	686	National education longitudinal study (NELS): second wave 1990
Stiefel et al. (2006)	Grade 5	USA (New York City)	P	Enrolment Subgroups: Asian (11 %) Black (36.1 %) Hispanic (37.4 %) White (15.5 %)	Citywide test in reading (CTB/McGraw Hill test of basic skills or New York State English language assessment)	958		Data set provided by the New York City department of education (2000–2001 school year)
	Grade 8		S	Subgroups: Asian (11.4 %) Black (36.3 %) Hispanic (34.6 %) White (17.7 %)		1221		
Sun et al. (2012)		Hong Kong	S	Total school enrolment (number of students) on 1 February 2006	Student science achievement (PISA 2006 science literacy test scores)	1039	174	PISA 2006 Hong Kong sample

(continued)

Table A.1 (continued)

Overview of studies of school size on student achievement

Authors	Sample	Country	School type[a]	School size measure	Student achievement measure	Mean	SD	Database
Tanner and West (2011)		USA (Georgia)	S	Net enrolment in the high school	Student achievement measured by 7 variables: Scholastic aptitude test (SAT) Graduation rate per school Averages scores on the Georgia high school graduation test (GHSGT): English Mathematics Science Social studies Writing	1370	682	Georgia department of education
Weiss et al. (2010)		USA (nationally representative)	S	Total number of students in school Categories: Small: 1–599 students (RF) Moderately small 600–999 Moderately large 1,000–1,599 Large 1,600–2,499	Mathematics achievement: standardized score derived from students' performance on the ELS:02 mathematics assessment			Educational longitudinal study of 2002 (ELS:2002) 10th grade
Wyse et al. (2010)		USA (nationally representative)	S	Total school enrolment Categories: 1–399 400–799 800–1,199 1,200–1,999 2,000 or more	Mathematics achievement: standardized score derived from students' performance on the ELS:2002 12th grade mathematics assessment			Educational longitudinal study of 2002 (ELS:2002) 10th and 12th grade

Table A.2 Methodological information available from studies of school size on student achievement

Methodological information available from studies of school size on student achievement											
Authors	Sample	Achievement measure	Number of schools	Number of classes/ teachers	Number of students	Statistical technique used	Value added	Effects reported in publication	SE reported	Direction of the effect	Further information
Åberg-Bengtsson (2004)		IEA reading literacy	124		3432	Multilevel SEM (standardized)	Yes	−0.02		n.s.	
Alspaugh (2004)		Composite (reading, math, language, science, social science)				ANCOVA <200 200–299 300–399 400–499 >= 500	Yes	ES = 0.471 ES = −0.016 ES = −0.391 ES = 0.107 ES = −0.169	0.169 0.105 0.182 0.258 0.076	Schools <200 highest mean composite score, category 300–399 lowest score	Three smallest enrolment groups of schools older inner-city schools, tow largest groups more newer suburban schools
Archibald (2006)		Reading Math	53	421	7,601	Multilevel (HLM) (standardized)	Yes	ß = −0.03 ß = −0.07	0.02 0.02	− −	s at 0.05 s at 0.05
Barnes et al. (2006)	KS1 KS2	KS1 English KS1 Math KS2 English KS2 Math KS2 Science				Regression	Yes			n.s. n.s. n.s. n.s. n.s.	
Bickel et al. (2001)		Reading Math Writing Size expressed in 1,000 student units Composite (ln) (size expressed in ln of single unit student)	1,001			Regression (unstandardized and standardized)	Yes	ß = 0.065 ß = 0.040 ß = 0.025 ß = 0.079		n.s. n.s. n.s. n.s.	*Siginificant negative effect from size-by-SES
Borland and Howsen (2003)		Combined subject scores for reading, language and mathematics	654	1,360	3,1440	2SLS Regression (unstandardized) School size School size[b]	Yes	ß = 0.40 *b* = −0.000	0.165 0.000	+ ∩ 760	
Bowles and Bosworth (2002)		Average school score for reading, writing and math	80			Regression (simultaneous equation modelling) unstandardized	Yes	*b* = −1.090		n.s.	

(continued)

Table A.2 (continued)

Methodological information available from studies of school size on student achievement

Authors	Sample	Achievement measure	Number of schools	Number of classes/ teachers	Number of students	Statistical technique used	Value added	Effects reported in publication	SE reported	Direction of the effect	Further information
Bradley and Taylor (1998)	11–16 schools 1992 1996 1992 1996	%> = 5 GCSEs at grades A* to C School size (pupils/100) School size (pupils/100) squared	1,307 1377			Ordered logistic regression (unstandardized)	Yes	$b = 0.55$ $b = 0.38$ $b = -0\ 0.024$ $b = -0\ 0.015$	0.0013 0.0011	+ + ∩ 1130 ∩ 1230	
	11–18 schools 1992 1996 1992 1996	%> = 5 GCSEs at grades A* to C School size School size (pupils/100) squared	1580 1514					$b = 0.0055$ $b = 0.0056$ $b = -0\ 0.021$ $b = -0\ 0.019$	0.0096 0.0098	+ + ∩ 1350 ∩ 1440	
Caldas (1993)	P S	Average school score composite	737 468			Regression (standardized)	Yes	ß = −0.06 ß = 0.043		+ n.s.	.
Carolan (2012)		12th grade Math score Small school Moderate Moderately large Large (RF)	579		9,647	Multilevel (unstandardized)	Yes	$b = -0.24$ $b = -0.68$ $b = -0.04$	0.40 0.36 0.31	n.s.	Indicator for moderately sized schools has a marginally significant relationship to 12th grade math scores. Students in these schools are predicted 0.04 sd less than students in large schools
Chen and Weikart (2008)		Mean school score English and math	212			Structural equation modelling (school level)	Yes	ß = −0.002		n.s.	The study presumed a direct effect from school disorder to achievement. An equally plausible relationship could be made that lower academic performance may lead to school disorder

(continued)

Table A.2 (continued)

Methodological information available from studies of school size on student achievement

Authors	Sample	Achievement measure	Number of schools	Number of classes/ teachers	Number of students	Statistical technique used	Value added	Effects reported in publication	SE reported	Direction of the effect	Further information
Coladarci and Cobb (1996)		Composite of English and math			4,567	Structural equation modelling (individual level)	Yes	ß = 0.005		n.s.	Compares smaller (<800) and larger (>= 1,600) schools
Deller and Rudnicki (1993)		Composite aggregated test score (reading, writing, math, science, social studies and humanities)	139			Regression (unstandardized)	Yes	$b = -0.040$		−	
Driscoll et al. (2003)	Primary Middle High school	Cailfornia academic performance index (composite)	4025 753 747			Regression (standardized)	Yes	ß = −0.073 ß = −0.0073 ß = 0.015		− n.s. n.s.	.
Durán-Narucki (2008)		% poor achievement English math	95			Regression (standardized)	Yes	ß = −0.017 ß = −0.088		n.s. n.s.	Mediation model (attendance is mediator).
Eberts et al. (1990)		Gain in math achievement Small schools versus medium schools Medium schools versus large schools	287		1,4000	Regression	Yes			−	The overall impact of size between medium and small schools is about 8 % of the typical gain in student achievement. The gain in large schools is 28 % lower than in medium schools
Fernandez (2011)		Aggregated growth in math reading math in reading	252			Regression (unstandardized) (standardized)	Yes	$b = 0.000$ $b = 0.000$ ß = 0.000 ß = 0.000	0.001 0.000 0.118 0.000	n.s. n.s.	
Foreman-Peck and Foreman-Peck (2006)		% of pupils in a school gaining 5 or more A-C GCSEs Ln school size Ln school size squared	1119			Logistic regression (unstandardized)	Yes	$b = 1.636$ $b = -0.129$		+ ∩ 560	s at 0.05 s at 0.05

(continued)

Table A.2 (continued)

Methodological information available from studies of school size on student achievement

Authors	Sample	Achievement measure	Number of schools	Number of classes/ teachers	Number of students	Statistical technique used	Value added	Effects reported in publication	SE reported	Direction of the effect	Further information
Fowler and Walberg (1991)		Minimum basic skills Test	293			Backward stepwise regression (unstandardized)	Yes	n.r.		n.s.	In the belief that some of the relationships may be curvilinear rather than linear explorations with quadratic terms were undertaken and no improvement was observed
		Average reading test score						n.r.		n.s.	
		Average math test score						$b = -0.001$		−	
		% passing reading test score						n.r.		n.s.	
		% passing math test score						n.r.		n.s.	
		% passing both reading and math score						n.r.		n.s	
								$b = -0.002$		−	
		High school proficiency test						n.r.		n.s	
		Average reading test score						$b = -0.001$		−	
		Average math test score						n.r.		n.s.	
		Average writing test score						$b = -0.002$		−	
		Average writing test mc score						$b = -0.005$		−	
		Average writing test essay score						n.r.		n.s.	
		% passing reading test score									
		% passing Math test score									
		% passing writing test									
Gardner et al. (2000).		(small schools vs. large schools)	127			An(c)ova	Yes	F(1.123) = 3.46		n.s.	
		Verbal SAT						F(1.123) = 18.79		+	
		Math SAT									
Heck (1993)		School math	235			Regression (standardized)	Yes	ß = −0.12		−	
		School reading						ß = −0.16		−	
Holas and Huston (2012)	Grade 5 Grade 6	Tested achievement Teacher-related achievement	10		855	Structural equation modelling	Yes	n.r.		n.s.	
		Teacher-related achievement						n.r.		n.s.	
								n.r.		n.s.	
Kahne et al. (2008)		Scores of the 11 grades on the Prairie state achievement exam reading scores	80			Multilevel (unstandardized)	Yes	Difference		n.s.	
		2002–2003						1.0		n.s.	
		2003–2004						−0.5		n.s.	
		2004–2005						−0.4		n.s.	
		2005–2006						−0.7			

(continued)

Table A.2 (continued)

Methodological information available from studies of school size on student achievement

Authors	Sample	Achievement measure	Number of schools	Number of classes/ teachers	Number of students	Statistical technique used	Value added	Effects reported in publication	SE reported	Direction of the effect	Further information
Kuziemko (2006)		Absolute change in average ISTEP math score change over 1 year 2 years 3 years Absolute change in average ISTEP language score change over 1 year 2 years 3 years	>100			Two SLS regression (unstandardized)	Yes	$b = -1.20$ $b = -3.841$ $b = -4.123$ $b = -0.2734$ $b = -1.187$ $b = -1.656$	2.09 1.427 2.250 1.115 1.029 1.765	n.s. – n.s. n.s. n.s. n.s.	
Lamdin (1995)		Teacher pupil ratio as input Reading Math Professional/pupil ratio as input Reading Math Expenditure per pupil as input Reading Math	97			Regression (unstandardized)	Yes	$b = -0.006$ $b = -0.001$ $b = -0.007$ $b = -0.002$ $b = -0.009$ $b = -0.003$		n.s. n.s. n.s. n.s. n.s. n.s.	
Lee and Loeb (2000)		1997 Math score Medium versus small Large versus small	264	4,495	2,2599	Multilevel (unstandardized)	Yes	Direct effects $b = -0.073$ $b = -0.041$		– n.s.	Direct and indirect effects: ES = −0.64 ES = −0.45
Lee and Smith (1995)		Gain in Mathematics Reading History Science	820		1,1794	Multilevel (standardized)	Yes	ES = −0.39 ES = −0.32 ES = −0.15 ES = −0.37		– – – –	

(continued)

Table A.2 (continued)

Methodological information available from studies of school size on student achievement

Authors	Sample	Achievement measure	Number of schools	Number of classes/ teachers	Number of students	Statistical technique used	Value added	Effects reported in publication	SE reported	Direction of the effect	Further information
Lee and Smith (1997)		Achievement gain math <300; 301–600; 601–900; 901–1,200; 1201–1,500; 1,501–1,800; 1,801–2,100; >2100 achievement gain reading <300; 301–600; 601–900; 901–1,200; 1,201–1,500 1,501–1,800; 1,801–2,100; >2,100	789		9,812	Multilevel (unstandardized)	Yes	$b = -0.931$ $b = -0.089$ $b = 1.512$ $b = 0.589$ b = −0.152 $b = -0.415$ $b = 1.842$ $b = -0.532$ $b = 0.149$ $b = 0.539$ $b = 0.290$ $b = -0.254$ $b = -0.455$ $b = -0.911$		∩ 601–900 ∩ 601–900	
Lubienski et al. (2008)	Grade 4	Math	6,288		157,161	Multilevel (unstandardized)	Yes	$b = 0.4$		n.s.	
	Grade 8	Math	4,870		119,364			$b = 0.6$		+	"School size slightly positively associated with achievement" (p. 129)
Luyten (1994)	USA 1st and 2nd sample	Math	58	104	2,212	Multilevel (unstandardized)	Yes	n.r.		n.s.	
	Sweden	Math	95	182	3,500			n.r.		n.s.	
	Netherlands math	Math	228	228	5,313			n.r.		n.s.	
	Netherlands science	Science	194	194	4,286			n.r.		n.s.	
Ma and McIntyre (2005)		Math Students/100	34		1,518	Multilevel		n.r.		n.s.	

(continued)

Table A.2 (continued)

Methodological information available from studies of school size on student achievement

Authors	Sample	Achievement measure	Number of schools	Number of classes/ teachers	Number of students	Statistical technique used	Value added	Effects reported in publication	SE reported	Direction of the effect	Further information
Maerten-Rivera et al. (2010)		Science	198		23854	Multilevel (unstandardized)	Yes	$b = -0.01$	0.01	n.s.	
McMillen (2004)	Primary school	Reading	1,053		54,615	Multilevel (unstandardized)	Yes				No significant main effect size reading or math achievement, but significant effect through interactions with students' prior level of achievement
		<400						$b = 0.05$	0.17	n.s.	
		400–549						$b = 0.12$	0.16	n.s.	
		550–699						$b = -0.05$	0.16		
		>700 (ref category)						$b = -0.26$	0.27		
		Math						$b = -0.09$	0.25		
		<400						$b = 0.18$	0.26		
		400–549									
		550–699									
		>700 (ref category)									
	Middle school	Reading	508		53,306	Multilevel (unstandardized)	Yes				No significant main effect size reading or math achievement, but significant effect through interactions with students' prior level of achievement
		<400						$b = -0.45$	0.21	n.s.	
		400–549						$b = -0.35$	0.20	n.s	
		550–699						$b = -0.46$	0.20		
		>700 (ref category)						$b = -0.91$	0.39		
		Math						$b = -0.69$	0.37		
		<400						$b = -0.89$	0.38		
		400–549									
		550–699									
		>700 (ref category)									
	High school	Reading	333		58,786	Multilevel (unstandardized)	Yes				In reading and math significant and positive main effect for size, along with statistically significant interactions involving size and ethnicity and size and parent education level (reading and math) and size and ethnicity (math)
		<700						$b = -2.58$	0.37	+	
		700–1,199						$b = -1.90$	0.33	+	
		1,200–1,699						$b = -0.91$	0.36		
		>1,700 (ref category)						$b = -5,14$	0.63		
		Math						$b = -3.93$	0.57		
		<700						$b = -2.64$	0.62		
		700–1,199									
		1200–1,699									
		>1,700 (ref category)									

(continued)

Table A.2 (continued)

Methodological information available from studies of school size on student achievement

Authors	Sample	Achievement measure	Number of schools	Number of classes/ teachers	Number of students	Statistical technique used	Value added	Effects reported in publication	SE reported	Direction of the effect	Further information
Moe (2009)	Primary		1,947			Regression (unstandardized)	Yes	$b = -11{,}05$ ES $= -0.16$	3.40	–	
	Secondary		829					$b = -6.16$ ES $= -0.20$	3.89	n.s.	
Ready and Lee (2006)		ECLS-K Literacy Kindergarten <275 276–400 401–600 -(RF) 601–800 >800 1st grade <275 276–400 401–600 (RF) 601–800 >800 ECLS-K Math Kindergarten <275 276–400 401–600 -(RF) 601–800 >800 1st grade <275 276–400 401–800 -(RF) 601–800 >800	527		7,740	Multilevel analysis (unstandardized)	Yes	$b = -0.04$ $b = 0.02$ $b = 0.02$ $b = -0.03$ $b = -0.03$ $b = 0.07$ $b = 0.04$ $b = -0.17$ $b = -0.03$ $b = 0.02$ $b = 0.00$ $b = -0.01$ $b = -0.13$ $b = 0.08$ $b = 0.02$ $b = -0.03$		n.s. n.s.	School size effects on learning in the lower elementary grades are distinctively nonlinear Learning in literacy (1st grade) is significantly disadvantaged in large schools Learning in mathematics (1st grade) is significantly advantaged in small schools

(continued)

Table A.2 (continued)

Methodological information available from studies of school size on student achievement

Authors	Sample	Achievement measure	Number of schools	Number of classes/ teachers	Number of students	Statistical technique used	Value added	Effects reported in publication	SE reported	Direction of the effect	Further information
Rumberger and Palardy (2005)		Achievement growth: test composite 1–600 (small) 601–1,200 RF 1,201–1,800 (large) >1,800 (extra large)	912		14,199	Multilevel analysis (standardized)	Yes	School effect sizes −0.000 0.124 0.105		∩ 1200–1,800	School effect size computed by first converting HLM coefficients to standard units and then dividing by the school-level standard deviation of the dependent variable estimated from the HLM null model
Sandy and Duncan (2010)	Urban	Composite Small schools (<1,000) (RF) Large schools (>1,000)			1,955	Regression (unstandardized)	Yes	$b = 1.667$		n.s.	
	Suburban	Composite Small schools (<1,000) (RF) Large schools (>1,000)						$b = 1.339$		n.s.	
Sawkins (2002)	1993–1994	Examination performance (Total number of pupils/100) (Total number of pupils/100) squared	398			Ordered logistic regression (unstandardized)	Yes	$b = -1.004$ $b = 0.042$		– U 1190	Only 4 % of schools were larger than the calculated minimum for 1993–1994
	1998–1999	Examination performance (Total number of pupils/100) (Total number of pupils/100) squared						$b = -0.565$ $b = 0.023$		– U 1230	Only 3.3 % of schools were larger than the calculated minimum for 1998–1999
Schneider et al. (2006)		Math 1–399 400–799 800–1,199 (RF) 1,200–1,999 2,000 or more	660		12,489	Multilevel analysis (unstandardized)	Yes	$b = -0.120$ $b = -0.287$ $b = -0.732$ $b = 0.288$	0.684 0.585 0.556 0.629	n.s.	

(continued)

Table A.2 (continued)

Methodological information available from studies of school size on student achievement

Authors	Sample	Achievement measure	Number of schools	Number of classes/ teachers	Number of students	Statistical technique used	Value added	Effects reported in publication	SE reported	Direction of the effect	Further information
Stewart, E.B. (2008)		Grade point average	715		11,999	Multilevel analysis (unstandardized)	Yes	$b = -0.05$	0.04	n.s.	
Stiefel et al. (2006)	Grade 5	Reading Subgroups: White (15.5 %) Black (36.1 %) Hispanic (37.4 %) Asian (11 %)	667		70,638	Multilevel analysis (unstandardized)	Yes	$b = -0.057$ $b = -0.037$ $b = -0.073$ $b = -0.052$	0.033 0.026 0.021 0.028	n.s. (all students)	Coefficients for white and Asian students at 0.10, coefficient for blank students n.s. Coefficient for Hispanics s at 0.01
	Grade 8	Reading Subgroups: White (17.7 %) Black (36.3 %) Hispanic (34.6 %) Asian (11.4 %)	278		55,921			$b = -0.064$ $b = -0.058$ $b = -0.034$ $b = -0.046$	0.045 0.030 0.028 0.046	n.s. (all students)	Coefficients for white, black, Hispanic and Asian students n.s.
Sun et al. (2012)		Science achievement	145		4,645	Multilevel analysis (unstandardized)	Yes	$b = 0.13$	0.02	+	
Tanner and West (2011)		SAT Graduation rate GHSGT: English Mathematics Science Social studies Writing	303			Regression	Yes	Change in R 0.001 0.000 0.003 0.001 0.002 0.001 0.035		n.s. n.s. n.s. n.s. n.s. n.s. s.	Direction of effect not reported
Weiss et al. (2010)		Math achievement Small 1–599 (RF) Moderately small 600–999 Moderately large 1,000–1,599 Large 1,600–2,499			10,946	Multilevel analysis (unstandardized)	Yes	$b = -0.821$ $b = -0.031$ $b = -0.020$	0.784 0.764 0.520	n.s.	

(continued)

Table A.2 (continued)

Methodological information available from studies of school size on student achievement

Authors	Sample	Achievement measure	Number of schools	Number of classes/ teachers	Number of students	Statistical technique used	Value added	Effects reported in publication	SE reported	Direction of the effect	Further information
Wyse et al. (2008)		Math achievement 1–399 matched with 2,000 or more 400–799 matched with 2,000 or more 800–1,199 matched with 2,000 or more 1,200–1,999 matched with 2,000 or more	745		12,853	WLS propensity score regression	Yes	Effect −0.565 −0.226 −0.031 −0.235		n.s.	Additional multivariate sensitivity analysis confirmed that there was not an particular school size that would results in optimal mathematics achievement

Notes

RF = reference category

n.s. = not significant at $p = 0.05$; *s* = significant

n.r. = not reported (not in table)

Table A.3 Results of vote counts examining the number of negative, nonsignificant, and positive effects of school size on academic achievement for each sample (school size measured as a continuous variable)

Study	Sample	School level	Negative	Not significant	Positive	Total
Archibald (2006)		P	2	0	0	2
Barnes et al. (2006)	KS1 7 years	P	0	2	0	2
	KS2 11 years	P	0	3	0	2
Bickel et al. (2001)		S	0	4	0	4
Borland and Howsen (2003)		P	0	0	1	1
Bowles and Bosworth (2006)		PS	0	1	0	1
Bradley and Taylor (1998)	11–16 1992	S	0	0	1	1
	11–16 1996	S	0	0	1	1
	11–18 1992	S	0	0	1	1
	11–18 1996	S	0	0	1	1
Caldas (1993)	Primary school	P	1	0	0	1
	Secondary school	S	0	1	0	1
Chen and Weikart (2008)		S	0	1	0	1
Coladarci and Cobb (1996)		S	0	1	0	1
Deller and Rudnicki (1993)		P	1	0	0	1
Driscoll et al. (2003)	Primary school	P	1	0	0	1
	Middle school	S	0	1	0	1
	High school	S	0	1	0	1
Durán-Narucki (2008)		P	0	2	0	2
Fernandez (2011)		PS	0	2	0	2
Foreman and Foreman-Peck (2006)		S	0	0	1	1
Fowler and Walberg (1991)		S	5	8	0	13
Heck (1993)		PS	2	0	0	2
Holas and Huston (2012)	Grade 5	P	0	2	0	2
	Grade 6	P	0	1	0	1

(continued)

Table A.3 (continued)

Study	Sample	School level	Negative	Not significant	Positive	Total
Maerten-Rivera et al. (2010)		P	0	1	0	1
Kahne et al. (2008)		S	0	4	0	4
Kuziemko (2006)		P	0	6	0	6
Lamdin (1995)		P	0	6	0	6
Lee and Smith (1995)		S	4	0	0	4
Lubienski et al. (2008)	Grade 4	P	0	1	0	1
	Grade 8	S	0	0	1	1
Luyten (1994)	USA	S	0	1	0	1
	Sweden	S	0	1	0	1
	Netherlands math	S	0	1	0	1
	Netherlands Science	S	0	1	0	1
Ma and McIntyre (2005)		S	0	1	0	1
Moe (2009)	Primary school	P	1	0	0	1
	Secondary school	S	0	1	0	1
Sawkins (2002)	1993–1994	S	1	0	0	1
	1998–1999	S	1	0	0	1
Stewart (2008)		S	0	1	0	1
Stiefel et al. (2006)	Grade 5	P	1	0	0	1
	Grade 8	S	0	1	0	1
Sun et al. (2012)		S	0	0	1	1
Tanner and West (2011)		S	0	6	0	6
Total			20	62	8	90

Table A.4 Results of vote counts examining the number of negative, nonsignificant, positive, and curvilinear effects of school size on academic achievement for each sample (school size effect modelled as quadratic function)

Study	Sample	School level	Direction of effect				Remarks
			−	n.s.	+	Curvilinear	
Borland and Howsen (2003)		P				∩ 760	Linear (+)
Bradley and Taylor (1998)	11–16 1992	S				∩ 1130	Linear (+)
	11–16 1996	S				∩ 1230	Linear (+)
	11–18 1992	S				∩ 1350	Linear (+)
	11–18 1996	S				∩ 1440	Linear (+)
Foreman-Peck and Foreman-Peck (2006)		S				∩ 560	Linear (+)
Sawkins (2002)	1993–1994	S				U 1190	Linear (−) Only 4 % of schools were larger than the calculated minimum for 1993–1994
	1998–1999	S				U 1230	Linear (−) Only 3.3 % of schools were larger than the calculated minimum for 1998–1999

Table A.5 Results of vote counts examining the number of negative, nonsignificant, positive, and curvilinear effects of school size on academic achievement for each sample (school size measured as categories)

Study	Sample	School level	Direction of effect				Remarks
			−	n.s.	+	Curvilinear	
Åberg-Bengtsson (2004)		P		1			Schools with an enrolment of less than 75 students and located in a rural district versus schools with an enrolment of 75 students and more
Alspaugh (2004)		P				<200 highest, 300–399 lowest	Categories: <200, 200–299, 300–399, 400–499, >= 500
Carolan (2012)		S		1			Categories: <600, 600–999, 1,000–1599, >1599 (RF)
Eberts et al. (1990)		P	1				<200 versus 400–599
			1				400–599 versus >800
Gardner et al. (2000)		S		1	1		200–600 versus >2,000
Inspectorate of Education (2003)		S		1			Categories: <500, 500–1000, >1000
Lee and Loeb (2000)		P	1	1			<400 versus 400–750
							400–750 versus >750
Lee and Smith (1997)		S				∩ 601–900 (2)	Categories: <300; 301–600; 601–900; 901–1,200; 1,201–1,500 (RF); 1,501–1,800; 1801–2,100; >2,100
McMillen (2004)	Primary school	P		2			Categories: <400, 400–549, 550–699, >700 (RF)
	Middle school	S		2			Categories: <400, 400–549, 550–699, >700 (RF)
	High school	S			2		Categories: <400, 400–549, 550–699, >700 (RF)

(continued)

Table A.5 (continued)

Study	Sample	School level	Direction of effect				Remarks
			−	n.s.	+	Curvilinear	
Ready and Lee (2006)		P		2		Significantly disadvantaged in large schools (2)	Categories: <275, 276–400, 401–600 (RF), 601–800, >800 Kindergarten, 1st grade
Rumberger and Palardy (2005)		S				∩ 1200–1800	Categories: 1–600, 601–1,200 (RF), 1,201–1,800, >1,800
Sandy and Duncan (2010)	Urban	S		1			Categories: <1,000 (RF), >1,000
	Suburban	S		1			
Schneider et al. (2007)		S		1			Categories: 1–399, 400–799, 800–1,199 (RF), 1,200–1,999, 2,000 or more
Weiss et al. (2010)		S		1			Categories: 1–599 (RF), 600–999, 1000-1599, 1600–2499
Wyse et al. (2008)		S		1			1–399 matched with 2,000 or more 400–799 matched with 2,000 or more 800–1,199 matched with 2,000 or more 1,200–1,999 matched with 2,000 or more
Total			3	16	3	6	

Students' and Teachers' Attitudes to School

See Tables A.6, A.7, A.8, A.9 and A.10.

Participation

See Tables A.11, A.12, A.13 and A.14.

Safety

See Tables A.15, A.16, A.17 and A.18.

Student Absence and Drop-Out

See Tables A.19, A.20, A.21, A.22, A.23, A.24, A.25 and A.26.

Other Student Outcomes

See Tables A.27, A.28, A.29 and A.30.

School Organization and Teaching and Learning

See Tables A.31, A.32, A.33 and A.34.

Costs

See Tables A.35, A.36 and A.37.

Table A.6 Overview of studies of school size on students' and teachers' attitudes to school

Overview of studies of school size on students' and teachers' attitudes to school								
Authors	Sample	Country	School type[a]	School size measure	Identification wit and connectedness to school measure	Mean	SD	Database
Bowen et al. (2000)		USA (nationally representative)	S	Total number of students enrolled in each school (middle school) Five categories: 0–399 400–599 600–799 800–999 1,000–1,399	School satisfaction (summary variable based on five items. e.g., I enjoy going to this school. I am getting a good education at this school) Teacher support (summary variable based on eight items. e.g., My teachers really care about e. I receive a lot of encouragement from my teachers)	689		Data collected by Louis Harris and Associates, Inc. (sample of 2,099 public schools in grades 6–12
Crosnoe et al. (2004)		USA (nationally representative)	S	School size: enrolment divided by 100 (School size)[b]: enrolment[b] divided by 100	Student school attachment: Adolescents reported the extent to which they agreed that in the past school year, they felt close to people at their schools, felt a part of their schools, and were happy to be at their schools. Student–teacher bonding (three items): the extent to which adolescents had trouble getting along with teachers, believed that teachers treated students fairly and, felt that teachers cared about them.	1381	838	National longitudinal study of adolescent health (in home sample 1995)
Holas and Huston (2012)	Grade 5	USA (nationally)	P	Total enrolment	School attachment: youth reported on the degree to which they felt competent in school, motivated and socially competent	490	210	NICHD study of early child care and youth development

(continued)

Table A.6 (continued)

Overview of studies of school size on students' and teachers' attitudes to school

Authors	Sample	Country	School type[a]	School size measure	Identification wit and connectedness to school measure	Mean	SD	Database
Kahne et al. (2008)		USA (Chicago)	S	School size	Teacher-teacher trust: the extent to which teachers believe that they have mutual respect for one another, for those who lead school improvement efforts and, for those who are experts at their craft. Academic personalism: students' views of their teachers' efforts to push students to higher levels of academic performance. Students also report on the degree to which they find their classes to be challenging. Classroom personalism: the degree to which students perceive that their teachers give individual attention to and are concerned about students' academic performance Sense of belonging: student reports of how personally connected students feel to the school Student–teacher trust: students' perceptions about the quality of their relationships with teachers Teacher support: students' reports of teachers' being there to help with personal matters			Consortium on Chicago school research's biannual survey. Administrative records of CPS and test data
Kirkpatrick Johnson et al. (2001)	Middle schools High schools	USA (nationally representative)	S	Total enrolment in schools in hundreds of students	School attachment: the extent to which adolescents agreed that. In the past school year, they felt close to people at their schools, felt a part of their schools. and were happy to be at their schools	1147	477 716	234	National longitudinal study of adolescent health
Koth et al. (2008)		USA (Maryland)	P	School enrolment	Achievement motivation (scale six items. e.g., My teachers believe I can do well in my school. I enjoy learning at this school)				Large scale study of a school-wide behavior support program called positive behavioral interventions and support

(continued)

Table A.6 (continued)

Overview of studies of school size on students' and teachers' attitudes to school

Authors	Sample	Country	School type[a]	School size measure	Identification wit and connectedness to school measure	Mean	SD	Database
Lee and Loeb (2000)		USA (Chicago)	P	Number of students in the school <400 (RF) (5.2 % of students) 400–750 (48.8 %) >750 (46.0 %)	Teachers' collective responsibility: the extent of a shared commitment among the faculty to improve the school so that all students learn. Teachers were asked how many of their colleagues feel responsible for student's academic and social development, set high standards of professional practice and take responsibility for school improvement			Data provided by the consortium on Chicago school research
McNeely (2002)		USA (nationwide)	S	Ln school size (in 100 s)	School connectedness (based on responses to five items: I feel close to people at this school, I feel like I am part of this school, I am happy to be at this school, the teachers at this school treat students fairly, I feel safe at my school)	642	765	National longitudinal study of adolescent health
Payne (2012)		USA (nationally representative)	S	Ln student enrolment	Communal school organization: composite of supportive and collaborative relations (between and among faculty and administration, support felt by teachers and views of relations between all school members) and common goals and norms (commonality of direction and expected behavior in the school)	792	479	Sample from national study of delinquency prevention
Rosenblatt (2001)		Israel (Northern part)	S	Number of students divided by 1,000	Organizational commitment (scale, 9 items), e.g., I am willing to put a great deal of effort beyond that normally expected in order to help this school be successful, I talk up this school to my friends as a great school to work in, I feel very little loyalty to this school	1020	650	Sample from secondary school teachers in the northern part of Israel

(continued)

Table A.6 (continued)

Overview of studies of school size on students' and teachers' attitudes to school

Authors	Sample	Country	School type[a]	School size measure	Identification wit and connectedness to school measure	Mean	SD	Database
Silins and Mulford (2004)		Australia	S	Size in 1997	Extent of students' engagement with school including students' perception of the way teachers relate to them, perceptions of their relationship with their peers, their perceptions of the usefulness of their schoolwork in later life, and the extent of identification with their school	632	283	Leadership for organizational learning and student outcomes (LOLSO)
Van der Vegt et al. (2005)		Netherlands	S	Number of pupils at school site	School connectedness (scale, eight items) Relationships with peers (scale, nine items) Relationship with teachers (scale, seven items)			National pupil monitor (secondary education) 2005
Vieno et al. (2005)		Italy (Veneto)	S	Size of the student body	Students' sense of community in the school: (scale, six items) I feel I belong at this school, Our students accept me as I am, Our school is a nice place to be, The students in my class enjoy being together, most of the students in my class are kind and helpful, when I need extra help I can get it from my teacher.	480	304	Health Behavior in School-aged Children (HBSC) project (Veneto regional data)
De Winter (2003)		Netherlands	S	School size <500 (32 %) 500–100 (28 %) >1,000 (40 %)	Classroom Climate			Health behavior of school children

Table A.7 Methodological information available from studies of school size on students' and teachers' attitudes to school

Methodological information available from studies of school size on students' and teachers' attitudes to school											
Authors	Sample	Identification wit and connectedness to school measure	Number of schools included	Number of classes/ teachers included	Number of students included	Statistical technique used	Value added	Effects reported in publication	SE reported	Direction of the effect	Further information
Bowen et al. (2000)		School satisfaction Teacher support	39		945	ANCOVA Least significant different pairwise comparisons	Yes			(>800) (>800)	Schools with enrolments of 800 or more may be too large to ensure a satisfactory educational environment
Crosnoe et al. (2004)		Student school attachment	84		13,162	Multilevel (unstandardized)[b] School size (100) School size (100) squared Multilevel (standardized) School size (100) School size (100) squared	Yes	$b = -0.02$ $b = 0.001$ $b = -0.0190$ $b = 0.011$	0.01 0.00 0.0095 0.00	– U 1900–2000	"…. students' positive views of their schools declined at a slowing rate as school size increased, with the lowest level occurring at schools with between 1,900–2,000 students." (p. 1268)
		Student–teacher bonding				Multilevel (unstandardized)[b] School size (100) School size (100) squared Multilevel (standardized) School size (100) School size (100) squared	Yes	$b = -0.02$ $b = 0.001$ $b = -0.226$ $b = 0.011$	0.01 0.00 0.113 0.00	n.s. n.s.	"The results for student–teacher bonding were similar to those for school attachment. …. with a steady decline bottoming out at about 1,700–1,800 students" (p. 1270) (significant at 0.10 level)
Holas and Huston (2012)	Grade 5	School attachment			827	Structural equation modelling	Yes	n.r.		n.s.	.
Kahne et al. (2008)		Teacher-teacher trust (teacher)	80			Multilevel (unstandardized)	Yes	$b = 0.209$		n.s.	
		Academic personalism (student)						$b = -0.523$		–	

(continued)

Table A.7 (continued)

Methodological information available from studies of school size on students' and teachers' attitudes to school

Authors	Sample	Identification wit and connectedness to school measure	Number of schools included	Number of classes/ teachers included	Number of students included	Statistical technique used	Value added	Effects reported in publication	SE reported	Direction of the effect	Further information
		Classroom (academic) personalism						$b = -0.523$		–	
		Sense of belonging (student)						$b = -0.175$		–	
		Student–teacher trust (student)						$b = -0.304$		–	
		Teacher support (student)						$b = -0.805$		–	
Kirkpatrick Johnson et al. (2001)	Middle schools High schools	School attachment	45 64		2,482 8,104	Multilevel (standardized)	Yes	n.r. n.r.		n.s. n.s.	
Koth et al. (2008)		Achievement motivation	37	120	2,468	Multilevel (unstandardized)	Yes	$b = -0.02$	0.01	–	
Lee and Loeb (2000)		Teachers' collective responsibility Medium versus small Large versus small	264	4,495	22,599	Multilevel (unstandardized)	Yes	$b = -0.406$ $b = -0.589$		– –	Teachers view's about the prevalence of collective responsibility, appeared to be more negative in medium sized schools (ES = −0.50) and even more in large schools (ES = −0.73) compared to small schools
McNeely (2002)		School connectedness	127		75,515	Multilevel (unstandardized		$b = -0.089$		–	
Payne (2012)		Communal school organization	253			OLS regression (standardized)	Yes	$\beta = -1.00$		n.s.	
Rosenblatt (2001)		Organizational commitment	12	200		Structural equation modelling	No	$\beta = 0.22$		–	
Silins and Mulford (2004)		Engagement	96	2,503	3,500	Structural equation modelling	Yes	$\beta = -0.16$		–	

(continued)

Table A.7 (continued)

Methodological information available from studies of school size on students' and teachers' attitudes to school

Authors	Sample	Identification wit and connectedness to school measure	Number of schools included	Number of classes/ teachers included	Number of students included	Statistical technique used	Value added	Effects reported in publication	SE reported	Direction of the effect	Further information
Van der Vegt et al. (2005)		School connectedness Relationship with peers Relationship with students	51		5,300	Regression (standardized)	No	n.r. ß = −0.06 ß = −0.06		n.s. − −	
Vieno et al. (2005)		Students' sense of community in the school	134	248	4,733	Multilevel (unstandardized)	Yes	$b = 0.001$	0.001	n.s.	
Winter (2003)		School climate < 500 (32 %) 500–1001(28 %) >1000 (40 %)			5,726	One way AN©OVA	Yes	Mean score 3.81 3.91 3.88		∩ 500–1000	Differences also significant after correction for school type or urbanisation

Notes [a] *n.s.* = not significant at $p = 0.05$

[b] standardized with s_x and s_y ß = bs_x/s_y

[c] *n.r.* = not reported (= not in table)

Table A.8 Results of vote counts examining the number of negative, nonsignificant, and positive effects of school size on students' and teachers' attitudes for each sample (school size measured as a continuous variable)

Study	Sample	Education level	Negative	Not significant	Positive	Total
Crosnoe et al. (2004)		S	1	1	0	2
Kahne et al. (2008)		S	4	1	0	5
Koth et al. (2008)		P	1	0	0	1
McNeely et al. (2002)		S	1	0	0	1
Payne (2012)		S	1	0	0	1
Rosenblatt (2001)		S	1	0	0	1
Silins and Mulford (2004)		S	1	0	0	1
Van der Vegt et al. (2005)		S	2	1	0	3
Vieno et al. (2005)		S	0	1	0	1
Total			12	4	0	16

Table A.9 Results of vote counts examining the number of negative, nonsignificant, positive, and curvilinear effects of school size on students' and teachers' attitudes for each sample (school size effect modelled as quadratic function)

Study	Sample	Education level	Direction of effect				Remarks
			−	n.s.	+	Curvilinear	
Crosnoe et al. (2004)		S				U 1900–2000	
				n.s.			

Table A.10 Results of vote counts examining the number of negative, nonsignificant, positive, and curvilinear effects of school size on students' and teachers' attitudes for each sample (school size measured as categories)

Study	Sample	Education level	Direction of effect				Remarks
			−	n.s.	+	Curvilinear	
Bowen et al. (2000)		S	−				Categories: 0–399, 400–599, 600–799, 800–999, 1,000–1,399
			−				Schools with enrolments of 800 or more may be too large to ensure a satisfactory educational environment
Lee and Loeb (2000)		P	−				Categories:
			−				<400 versus 400–750
							400–750 versus >750
Weiss et al. (2010)		S	−				Categories: 1–599 (RF), 600–999, 1,000–1,599, 1,600–2,499
Winter (2003)		S				1	Categories: <500, 500–1,000, >1,000

Table A.11 Overview of studies of school size on participation (students, teachers and parents)

Overview of studies of school size on participation (students, teachers and parents)								
Authors	Sample	Country	School type[a]	School size measure	Involvement measure	Mean	SD	Database
Coladarci and Cobb (1996)		USA (nationally representative)	S	Total high school enrolment Compares smaller (<800) and larger (>=1,600) schools Students from schools with other sizes eliminated from analyses	Extracurricular participation: self-reported participation for specific activities related to academics (e.g., clubs. student government. sport and the performing arts). Total EP composite across all activities and both grades (TEP)			National Education Longitudinal Study of 1988 (NELS: 88)
Crosnoe et al. (2004)		USA (nationally representative)	S	School size: enrolment divided by 100 (School size)[b]: (Enrolment divided by 100) squared	Student extracurricular participation Students reported whether they had engaged in 33 extracurricular activities. Activities are grouped into five categories (athletic, academic, performing arts, leadership, and other) and then summed	1381	838	National Longitudinal Study of Adolescent Health (in home sample 1995)
Dee et al. (2007)		USA (nationally representative)	S	School size Five categories: <400 (4.3 %) (RF) 400–799 (12.5 %) 800–1,199 (17.6 %) 1,200–2,199 (42.0 %) >2,200 (23.6 %)	Parent(s) act as a volunteer at the school (individual level)			Educational longitudinal study of 2002 (ELS: 2002)

(continued)

Table A.11 (continued)

Overview of studies of school size on participation (students, teachers and parents)

Authors	Sample	Country	School type[a]	School size measure	Involvement measure	Mean	SD	Database
Feldman and Matjasko (2006)		USA (nationally representative)	S	School size Three categories: 1–400 (18.7 %) (RF) 401–1,000 (47.1 %) >1,000 (34.2 %)	Adolescent school-based extracurricular activity participation (multiple activities)			National longitudinal study of adolescent health (Wave 1)
Gardner et al. (2000)		USA (California)	S	School size Small schools (200–600 pupils) Large schools (>2,000 pupils)	Average of total number of California parent teacher association members for each affiliated school for the years 1995–1996 en 1996–1997	424 2500		Data obtained through the California parent teacher association, or from questionnaires mailed to principals and/or by telephone contact
Holas and Huston (2012)	Grade 6	USA (nationally)	PS	Total enrolment	School involvement	690	300	NICHD study of early child care and youth development
Kahne et al. (2008)		USA (Chicago)	S	School size	Teacher influence: measures the extent of teachers' involvement in school decision making			Consortium on Chicago School Research's biannual survey. Administrative records of CPS and test data
Lay (2007)		USA (nationwide)	S	School size Continuous measure Categories based on continuous measure <300, 301–600, 601–900, 901–1200, 1201–1500, 1501–1800, >1800 Categories based on parental responses <300, 300–599, 600–999, >1000	Participation in school activities such as sports teams. safety patrol or school clubs			National household education survey

(continued)

Table A.11 (continued)

Overview of studies of school size on participation (students, teachers and parents)

Authors	Sample	Country	School type[a]	School size measure	Involvement measure	Mean	SD	Database
MacNeal (2008)		USA (nationwide, public schools)	S	Number of students (natural logarithm)	Number of school related activities in which a student participated (subject matter clubs, band and orchestra, chorus and dance, athletic teams, cheerleading, pep clubs and majorettes, vocational education clubs and hobby clubs) (scale 0–8)	1053		High school and beyond database (NCES 1983)
Silins and Mulford (2004)		Australia	S	School size in 1997	Participation—representing the extent of students' participation in school including absences, participation in extracurricular activities, preparedness to do extra schoolwork, involvement in classroom/school decisions and setting own learning goals, and voicing opinion in class	632	283	Leadership for organizational learning and student outcomes (LOLSO)

Table A.12 Methodological information available from studies of school size on participation (students, teachers and parents)

Methodological information available from studies of school size on participation (students, teachers and parents)

Authors	Sample	Involvement measure	Number of schools included	Number of classes/ teachers included	Number of students included	Statistical technique used	Value added	Effects reported	SE reported	Direction of the effect	Further information
Coladarci and Cobb (1996)		Total extracurricular participation (individual level)			4,567	Structural equation modelling	Yes	ß = −0.210		−	Compares smaller (<800) and larger (>= 1,600) schools
Crosnoe et al. (2004)		Student Extracurricular participation	84		1,3420	Multilevel (unstandardized)[b] School size (100) School size (100) squared Multilevel (standardized) School size (100)	Yes	*b* = −0.02 n.r *b* = −0.0075	0.00 0.00	−	Extracurricular participation … school size had a negative linear effect (p. 1270)
Dee et al. (2006)		Parent(s) act as a volunteer	390		8,197	Multilevel (unstandardized)[b] School size <400 (RF) 400–799 800–1199 1200–2199 >2200 Multilevel (standardized) School size <400 400-799 800-1199 1200-2199 >2200	Yes	*b* = −0.086 *b* = −0.080 *b* = −0.108 *b* = −0.123 ß = −0.066 ß = −0.070 ß = −0.123 ß = −0.120	0.028 0.027 0.027 0.030 0.021 0.024 0.031 0.030	–(n.s.)	"…, it should be noted the estimates for most categorical school size indicators above 400 students are not statistically distinguishable from each other" (p. 15)
Feldman and Matjasko (2006)		Extracurricular activity participation (multiple activities)	132		13,810	Multinomial logistic regression 1–400(RF) 401–1,000 >1,000	Yes	Relative risk ratio 0.63 0.39	0.13 0.09	−	Adolescents from medium and large schools (compared with small) were more likely to be nonparticipants than to participate in multiple activities. Difference between small and medium schools n.s.

(continued)

Table A.12 (continued)

Methodological information available from studies of school size on participation (students, teachers and parents)

Authors	Sample	Involvement measure	Number of schools included	Number of classes/ teachers included	Number of students included	Statistical technique used	Value added	Effects reported	SE reported	Direction of the effect	Further information
Gardner et al. (2000)		Parental participation (small vs. large schools)	127			An(c)ova	Yes	F(1.118) = 13		–	
Holas and Huston (2012)	Grade 6	School involvement			825	Structural equation modelling	Yes	ß = −0.13		–	
Kahne et al. (2008).		Teacher influence (teacher)	80			Multilevel (unstandardized)	Yes	$b = -0.871$		–	
Lay (2007)		Participation in school activities			3,010	Logistic regression	Yes	$b = -0$	0.00004	–	Participation in schools with 1,501–1,800 students significantly less likely
								0.0001	0.186	n.s.	
		School size as continuous measure						$b = 0.322$	0.155		
								$b = 0.119$	0.153		No relationship between school size and participation
								$b = 0.198$	0.151		
		Categories based on continuous measure						$b = -0.073$	0.144		
		<300						$b = -0.287$	0.125		
		301–600						$b = -0.189$	0.168		
		601–900						$b = -0.253$	0.117		
		901–1200						$b = -0.110$	0.100		
		1501–1800						$b = -0.100$			
		>1800									
		Categories based on parental responses									
		<300									
		600–999									
		>1,000									

(continued)

Table A.12 (continued)

Methodological information available from studies of school size on participation (students, teachers and parents)

Authors	Sample	Involvement measure	Number of schools included	Number of classes/ teachers included	Number of students included	Statistical technique used	Value added	Effects reported	SE reported	Direction of the effect	Further information
McNeal (2002)		Participation in: school related activities Athletics	281		5,772	Multilevel (unstandardized) Hierarchical logistic regression	Yes	$b = -0.309$ $b = -0.445$		– –	
Silins and Mulford (2004)		Participation	96	2,503	3,500	Structural equation modelling	Yes	ß = −0.39		–	

* = included in vote count

Notes [a] *n.s.* = not significant at $p = 0.05$, [b] standardized with s_x and s_y ß = $bs_x/s_{y,}$ [c] *n.r.* = not reported (= not in table)

Table A.13 Results of vote counts examining the number of negative, nonsignificant, and positive effects of school size on participation for each sample (school size measured as a continuous variable)

Study	Sample	Education level	Negative	Not significant	Positive	Total
Coladarci and Cobb (1996)		S	1	0	0	1
Crosnoe et al. (2004)		S	1	0	0	1
Kahne et al. (2008)		S	1	0	0	1
Kirkpatrick Johnson et al. (2001)	Grade 7–11 Middle school	S	0	1	0	1
	Grade 7–11 High school	S	0	1	0	1
Lay (2007)		S	1	0	0	1
MacNeal (2008)		S	2	0	0	2
Silins and Mulford (2004)		S	1	0	0	1
Total			7	2	0	9

Table A.14 Results of vote counts examining the number of negative, nonsignificant, positive, and curvilinear effects of school size on participation for each sample (school size measured as categories)

Study	Sample	Education level	Direction of effect				Remarks
			−	n.s.	+	Curvilinear	
Dee et al. (2007)		S		n.s.			Categories: <400 (RF), 400–799, 800–1199, 1200–2199, >2200 "…, it should be noted the estimates for most categorical school size indicators above 400 students are not statistically distinguishable from each other" (p. 15)
Feldman and Matjasko (2006)		S	−				Categories: 1–400, 401–1000, >1000 Adolescents from medium and large schools (compared with small) were more likely to be nonparticipants than to participate in multiple activities. Difference between small and medium schools n.s.
Gardner et al. (2000)		S	−				Small schools (200–600) versus large schools (>2000)
Lay (2007)		S				Participation in schools with 1,501–1,800 students significantly less likely	Categories based on continuous measure <300, 301–600, 601–900, 901–1200, 1201–1500, 1501–1800, >1800
		S		n.s.			Categories based on parental responses <300, 300–599, 600–999, >1000

Table A.15 Overview of studies of school size on safety

Overview of studies of school size on safety

Authors	Sample	Country	School type[a]	School size measure	Safety measure	Mean	SD	Database
Attar-Schwartz (2009)		Israel	S	Number of students at school	Sexual harassment victimization (0 = never sexually harassed in the past month … 7 = victimized by seven different sexual harassment behaviors) (school average)	557	332	National survey of school violence among students in grade 4 through 11 in Israel during spring 2005
Bonnet et al. (2009)		Netherlands (Flevoland and North Holland provinces)	P	Number of children attending the school Small schools (<300 pupils) (7 schools) Medium schools (301–500 pupils) (13 schools) Large schools (>500 pupils) (3 schools)	Victimization score of individual child (16 items. teacher questionnaire)			2003 children in the 1st two grades of elementary schools
Bowen et al. (2000)		USA (nationally representative)	S	Total number of students enrolled in each school (middle school) Five categories: 0–399 (6 schools) 400–599 (11 schools) 600–799 (8 schools) 800–999 (10 schools) 1000–1399 (4 schools)	School safety: summary variable ranging from 10 to 30. Scale (10 items) assessing whether various student behaviors were a big problem, a slight problem or no problem at the school with regard to fights among students, students use of alcohol, student physical and verbal abuse of teachers	689		Data collected by Louis Harris and Associates, Inc. (sample of 2099 public schools in grades 6–12
Bowes et al. (2009)		England (twins register)	P	Total number of children in school based on data for schools attended by study participants	Involvement in bullying between ages 5 and 7 Groups: Noninvolved (RF) Victims (victimized by bullies but have not bullied others). Bullies have bullied others but have not themselves been victimized) Bully-victims (have been victimized by bullies and have bullies others as well)	291	136	Environmental risk longitudinal twin study

(continued)

Table A.15 (continued)

Overview of studies of school size on safety

Authors	Sample	Country	School type[a]	School size measure	Safety measure	Mean	SD	Database
Chen (2008)		USA (nationally representative)	S	Enrolment Less than 300 (7.2 %) 300–499 (12.1 %) 500–999 (26.1 %) 1,000 or more (54.6 %)	Number of crime incidents that occurred in a school in the past 12 months Misbehavir (frequency of student bullying and frequency classroom disorder)			2,000 school survey on crime and safety
Chen and Weikart (2008)		USA (New York City)	S	Number of students enrolled at each school	School disorder: major crime, minor crime and noncrime incidents reported by New York police department on a per 1,000-student population basis (NYPD reside on campus and are responsible for school safety)	960	493	Data from New York City department of education (2002–2003 and 2003–2004 school year data for all middle schools)
Chen and Vazsonyi (2013)		USA (nationally representative)	S	Small schools (<400 students) (14 %) Medium size schools (400–1,000 students) (38 %) Large schools (>1,000 students) (48 %)	Problem Behavior Health survey (Participant scale): 17 items range of norm-violating behaviors in the past 12 months, alcohol use and marijuana			First two waves of the national longitudinal study of adolescent health (add health)
Eccles et al. (1991)		USA (nationally representative)	P and S	Total school enrolment P/K/1–8 th grade schools versus middle and junior high schools (grades 6–8, grades 7–8, grades 7–9) P/K/1–8 th grade schools smaller on the average than other three types	Violence (student and teacher questionnaire, e.g., physical conflicts between students is not a problem. student possession of weapons is not a problem) Substance abuse while at school (student and teacher questionnaire, e.g., student use of alcohol is not a problem, student use of illegal drugs is not a problem)			National education longitudinal study (NELS: 88)

(continued)

Table A.15 (continued)

Overview of studies of school size on safety

Authors	Sample	Country	School type[a]	School size measure	Safety measure	Mean	SD	Database
Gottfredson and DiPietro (2011)		USA (nationally representative)	S	Number of students enrolled in the school (natural log)	Personal victimization: five items students' in school personal victimization experiences (e.g., been physically attacked, been threatened with a knife) Property victimization: two items students' in school personal victimization experiences (e.g., whether the respondent has something stolen from his or her desk.)	792	478	Sample drawn from most comprehensive list of schools available
Haller (1992)		USA (nationally representative)	PS	Enrolment in 1980	Disorder (reported by principals): a score on the seriousness of five types of disciplinary problems in their school (theft, vandalism, drugs, weapons and verbal abuse of teachers) Disorder (reported by students): % of students who perceived four types of disorder to be a problem in their school Disorder (self-reported: % of students that had personally been a discipline problem. or had been suspended)	963	1219	High School and Beyond surveys (1980 en 1982)
Heck (1993)		USA (Western state)	PS	Actual size of enrolment	Substance abuse: the number of student suspensions for significant offences (i.e. felonies and misdemeanours)			State department of education's survey on restructuring the curriculum

(continued)

Table A.15 (continued)

Overview of studies of school size on safety

Authors	Sample	Country	School type[a]	School size measure	Safety measure	Mean	SD	Database
Inspectorate of Education (2003)		Netherlands	S	School size Categories: <500 501–1,000 >1,000	Pupil guidance/support School climate			Data form regular school supervision (years 1999–2000, 2000–2001)
Kahne et al. (2008)		USA (Chicago)	S	School size	Respectful classroom behavior: students' assessment of their peers' classroom behavior with regard to how they treat one another, how often they disrupt class, if they have respect for one another and, if they help one another learn			Consortium on Chicago School Research's biannual survey. Administrative records of CPS and test data
Khoury-KassabrI et al. (2004)		Israel (nationally representative)	S	Number of students attending the school	Physical victimization, including being cut with a knife and getting a beating that resulted in medical treatment. Victimization by threats (e.g., being threatened verbally or with a knife) Victimization by moderate physical violence: three items indicating whether students were kicked or punched Verbal–social victimization: examples include being mocked, insulted or humiliated	505	298	National school violence survey carried out across Israel during the spring of 1999

(continued)

Table A.15 (continued)

Overview of studies of school size on safety

Authors	Sample	Country	School type[a]	School size measure	Safety measure	Mean	SD	Database
Klein and Cornell (2010)		USA (Virginia)	S	School enrolment size	Self-report bully victimization Student perceptions of bullying Teacher perceptions of bullying Total bullying violations Bullying violations rate Self-report threat victimization Total threat violations Threat violations rate Self-report physical attack victimization Total attack violations Attack violations rate	1210	690	Virginia high school safety study
Koth et al. (2008)		USA (Maryland)	P	School enrolment	Order and discipline (scale 11 items): e.g., my school is a safe place. Children in my school fight a lot			Large scale study of a school-wide behavior support program called positive behavioral interventions and support
Leung and Ferris (2008)		Canada (Montreal, Quebec)	S	The actual number of students (in thousands) in the school that the subject attended Categories <999 (RF) 1,000–1,499 1,500–1,999 >2,000	If the subject self-reported that at least one violent event took place when 17 years old. Violent events are defined as participation in a gang fight, fighting with fists, using a weapon in a fight, carrying weapons, beating up someone without reason, beating up someone to force that person to do something, destroying something that belongs to a family member			Data collected from a young group of males, their families and peers form an ongoing longitudinal youth study. Data on school sizes and other school characteristics collected from studies in education policy

(continued)

Table A.15 (continued)

Overview of studies of school size on safety

Authors	Sample	Country	School type[a]	School size measure	Safety measure	Mean	SD	Database
Mooij et al. (2011)		Netherlands	S	Number of pupils/100	Pupil's feelings of safety at school	926	514	National survey about school safety in secondary education (initiated by Dutch ministry of education, culture and science) (2nd date wave, 2008)
O'Moore et al. (1997)	P	Ireland	P	Number of pupils 0–199 (small) 200–499 (medium) 500 or more (large)	Incidence of being bullied Incidence of bullying others			
	S		S	0–199 (small) 200–499 (medium) 500 or more (large)				
Stewart (2003)		USA (nationally representative)	S	School enrolment	School misbehavior (scale, four items): got in trouble for not following school rules, put on in-school suspension, Suspended or put on probation from school, Got into a physical fight at school)	1540	686	National education longitudinal study (NELS): second wave 1990
Van der Vegt et al. (2005)		Netherlands	S	Number of pupils at school site	Safety (feelings) Safety policy Bullying and fighting Vandalism, drugs and theft			National pupil monitor (secondary education) 2005
Watt (2003)	Males Females	USA (nationally representative)	S	Three categories: Small (400 or less students) (19.2 %) Medium (401–1000) (46.7 %) Large (1001–4000) (34.2 %)	Violence: weapon use/threat How often in the 12 past months did you use or threaten to use a weapon to get something from someone? Dichotomy to differentiate those who had used or threaten to use a weapon in the past year from those who had not			National longitudinal survey of adolescent health (add health) Panel study conducted in 1994–1996 Wave 1 and 2 Data analyzed separately for males and females

(continued)

Table A.15 (continued)

Overview of studies of school size on safety

Authors	Sample	Country	School type[a]	School size measure	Safety measure	Mean	SD	Database
Wei et al. (2010)		Taiwan (Taichung City)	S	Total number of students in school	Physical bullying (3 items): you hit other students, You kick other students, you hurt other students with dangerous objects or tools. Verbal bullying (two items): you verbally insult other students, You threaten other students.	1568	989	Public middle school subsample of the Taichung City youth life conditions survey
Winter (2003)		Netherlands	S	School size <500 (32 %) 500–1002 (28 %) >1000 (40 %)	Being bullied Bullying Frequent fighting			Health behavior of school children

Table A.16 Methodological information available from studies of school size on safety

Methodological information available from studies of school size on safety

Authors	Sample	Safety measure	Number of schools	Number of classes	Number of students	Statistical technique used	Value added	Effects reported	SE reported	Direction effect	Further information
Attar-Schwartz (2009)		Sexual harassment	327		16,604	Multilevel (HLM) (unstandardized)[b]	Yes	ß = 0.032	0.015	–	*s* at 0.05
Bonnet et al. (2009)		Victimization score	23	98	2,003	Multilevel analysis (unstandardized)	Yes	Small size b = −0.17 Medium size b = −0.16 Large size (RF) b = 0.165	0.09 0.07 0.08	(−) n.s.	
Bowen et al. (2000)		School safety	39		945	ANCOVA Least significant different pairwise comparisons	Yes	Schools with enrolments of 800 or more may be too large to ensure a satisfactory educational environment		(>800)	
Bowes et al. (2009)		Noninvolved (RF) Victims Bullies Bully-Victims			Sample of 2,232 children	Multivariate multinomial logistic regression	Yes	Odds ratio 1.2 0.9 0.8		– n.s. n.s.	School size is associated with an increased risk for being a victim of bullying
Chen (2008)		Misbehavior (student bullying and classroom disorder)				Structural equation modelling	Yes	ß = 0.12		–	
		Number of crime incidents						ß = 0.31		–	
Chen and Weikart (2008)		School disorder: Major crime, minor crime and noncrime incidents	213			Structural equation modelling (school level)	Yes	ß = −0.10		n.s.	
Chen and Vazsonyi (2013)		Problem behavior Small school (RF) Medium-size school Large school	85		9,163	Multilevel analysis (unstandardized)	Yes	*b* = 0.118 *b* = 0.172	0.073 0.076	–	

(continued)

Table A.16 (continued)

Methodological information available from studies of school size on safety

Authors	Sample	Safety measure	Number of schools	Number of classes	Number of students	Statistical technique used	Value added	Effects reported	SE reported	Direction effect	Further information
Eccles et al. (1991)		Violence	729			Regression	No	ß = −0.047		−	"the larger the school size, … the more violence were reported as being a major problem at the school by both teachers and students" (p. 351)
Eccles et al. (1991)		Substance abuse	759					ß = −0.059		−	"the larger the school size, … substance abuse were reported as being a major problem at the school by both teachers and students" (p. 351)
Gottfredson and DiPietro (2011)		Personal victimization Property victimization	253		13,597	Multilevel analysis (unstandardized)	Yes	$b = -0.005$ $b = -0.036$	0.003 0.011	n.s. −	
Haller (1992)		Disorder Principals Students Self-reported	558			Regression (standardized)	Yes	ß = 0.263 ß = 0.079 ß = 0.128		− n.s. −	
Heck (1993)		Suspensions	235			Regression (standardized)	Yes	ß = −0.03		n.s.	.
Inspectorate of Education (2003)		Pupil guidance and school climate	378			ANOVA	No			n.s.	
Kahne et al. (2008)		Respectful classroom behavior	80			Multilevel analysis (unstandardized	Yes	$b = -0.115$		n.s.	

(continued)

Table A.16 (continued)

Methodological information available from studies of school size on safety

Authors	Sample	Safety measure	Number of schools	Number of classes	Number of students	Statistical technique used	Value added	Effects reported	SE reported	Direction effect	Further information
Khoury-KassabrI et al. (2004)		Serious physical victimization	162		10,400	Multilevel analysis (unstandardized)	Yes	$b = 0.005$		n.s.	.
		Threats						$b = 0.007$		n.s.	
		Moderate physical victimization						$b = 0.002$		n.s.	
		Verbal-social victimization						$b = 0.007$		n.s	
Klein and Cornell (2010)		Self-report bully victimization	290		7,431	Regression (standardized)	Yes	ß = −0.01		n.s.	
		Student perceptions of bullying						ß = 0.41		−	
		Teacher perceptions of bullying						ß = 0.20		−	
		Total bullying violations						ß = 0.49		−	
		Bullying violations rate						ß = −0.33		+	
		Self-report threat victimization						ß = 0.06		−	
		Total threat violations						ß = 0.53		−	
		Threat violations rate						ß = −0.39		+	
		Self-report physical attack victimization						ß = −0.02		n.s.	
		Total attack violations						ß = 0.67		−	
		Attack violations rate						ß = −0.46		+	

(continued)

Table A.16 (continued)

Methodological information available from studies of school size on safety

Authors	Sample	Safety measure	Number of schools	Number of classes	Number of students	Statistical technique used	Value added	Effects reported	SE reported	Direction effect	Further information
Koth et al. (2008)		Order and discipline	37	120	2468	Multilevel analysis (unstandardized)	Yes	$b = 0.0$	0.01	n.s.	
Leung and Ferris (2008)		Youth violence (Actual number of students in 1000)	110		616	Logistic regression (unstandardized)	Yes	$b = 0.3475$		−	"In terms of its marginal effect, an increase in school enrolment would lead to about a 10 % increase in the probability of teenage violence" (p. 328)
		Size as continuous measure				Coefficient		0.0860			
						Marginal effect					
		Four categories:				Coefficient					Discontinuity in the effect of school size: "… teenagers who attended a school with more than 2,000 students were about 22 per cent more likely to engage in violent behavior than those who attended schools with less than 1,000 students" (p. 328)
		<999 (RF)								n.s.	
		1,000–1,499				Marginal effect		$b = 0.2196$		n.s.	
		1,500–1,999						$b = 0.8838$		−	
		>2,000						0.0680			
		<999 (RF)						0.0543			
		1,000–1,499						0.2186			
		1,500–1,999									
		>2,000									
Mooij et al. (2011)		Pupil's feelings of safety at school	104		26,162	Multilevel analysis (unstandardized)[b]	Yes	ß = 0.035	0.016	+	
O'Moore et al. (1997)	P	Incidence of being bullied	320		9,559	ANOVA	No			n.s.	No significant differences
		Incidence of bullying others								U	Highest proportion of pupils bullying others in medium-sized schools, least in large schools

(continued)

Table A.16 (continued)

Methodological information available from studies of school size on safety

Authors	Sample	Safety measure	Number of schools	Number of classes	Number of students	Statistical technique used	Value added	Effects reported	SE reported	Direction effect	Further information
	S	Incidence of being bullied Incidence of bullying others	211		10,843	ANOVA	No			U +	Incidence of pupils being bullied least common in large schools Being bullied least common in large-sized schools. Small schools highest incidence of bullying
Stewart (2003)		School misbehavior+	528		10,578	Nonlinear hierarchical generalized linear model (HGLM)	Yes	$b = 0.173$	0.042	−	After a wide range of individual and school level covariates is controlled for, school misbehavior continued to vary significantly between schools
Van der Vegt et al. (2005)		Safety (feelings) Safety policy+ Bullying and fighting+ Vandalism, drugs and theft+	51		5,300	Regression (standardized)	No	n.r. ß = −0.08 ß = 0.07 ß = 0.16		n.s. + − −	
Watt (2003)	Males Females	Weapon use/ threat Males Small school Medium school Females Small school Medium school			12,150	Logistic regression (unstandardized)	Yes	$b = 0.236$ $b = -0.146$ $b = 0.623$ $b = 0.183$	0.311 0.210 0.355 0.307	Total n.s.	Odds ratio 1.266 0.864 1.865 1.201 Combined vote count for males and females n.s.

(continued)

Table A.16 (continued)

Methodological information available from studies of school size on safety

Authors	Sample	Safety measure	Number of schools	Number of classes	Number of students	Statistical technique used	Value added	Effects reported	SE reported	Direction effect	Further information
Wei et al. (2010)		Physical bullying	12	36	1,172	Multilevel analysis (unstandardized)[b]	No	$b = -0.000$	0.000	n.s.	Intra-Class Correlation
		Verbal bullying						$b = -0.000$	0.000	n.s.	Physical bullying 0.03
		Physical bullying				standardized		ß = −0.073	0.116		Verbal bullying 0.02
		Verbal bullying						ß = −0.148	0.093		
Winter (2003)		Being bullied+			5,726	One way AN©OVA	Yes	<500 500–1000 >1000		+	Also significant after
		Bullying+						1.57 1.42 1.37		+	orrection for school
		Frequent fighting+						1.73 1.57 1.51		+	type or urbanisation
								1.87 1.64 1.67			

* = included in vote-count

Notes [a] *n.s.* = not significant at $p = 0.05$, [b] standardized with s_x and s_y ß = bs_x/s_y, [c] *n.r.* = not reported (= not in table)

Table A.17 Results of vote counts examining the number of negative, nonsignificant, and positive effects of school size on safety for each sample (school size measured as a continuous variable)

Study	Sample	Education level	Negative	Not significant	Positive	Total
Attar-Schwartz (2009)		S	1	0	0	1
Bowes et al. (2009)		P	1	2	0	3
Chen (2008)		S	2	0	0	2
Chen and Weikart (2008)		S	0	1	0	1
Eccles et al. (1991)		PS	2	0	0	2
Gottfredson and DiPietro (2011)		S	1	1	0	2
Haller (1992)		S	2	1	0	3
Heck (1993)		PS	0	1	0	1
Kahne et al. (2008)		S	0	1	0	1
Khoury-Kassabri et al. (2004)		S	0	4	0	4
Klein and Cornell (2010)		S	5	3	3	11
Koth et al. (2008)		P	1	1	0	0
Leung and Ferris (2008)		S	1	0	0	1
Mooij et al. (2011)		S	0	0	1	1
Stewart (2003)		S	1	0	0	1
Van der Vegt et al. (2005)		S	2	1	1	4
Wei et al. (2010)		S	0	2	0	2
Total			19	17	5	40

Table A.18 Results of vote counts examining the number of negative, nonsignificant, positive, and curvilinear effects s of school size on safety for each sample (school size measured as categories)

Study	Sample	Education level	Direction of effect				Remarks
			−	n.s.	+	Curvilinear	
Bonnet et al. (2009)		P		n.s.			Categories: <300, 301–500, >500
Bowen et al. (2000)		S	−				Categories: 0–399, 400–599, 600–799, 800–999, 1000–1399 Schools with enrolments of 800 or more may be too large to ensure a satisfactory educational environment
Chen and Vazsonyi (2013)		S	−				Categories: <400, 400–1000, >1000
Inspectorate of Education (2003)		S		n.s.			Categories: 0-500, 500–1000, >1000
Leung and Ferris (2008)		S		n.s.			Categories: <999 (RF), 1000–1499, 1500–1999, >2000 Discontinuity in the effect of school size: "… teenagers who attended a school with more than 2,000 students were about 22 % more likely to engage in violent behavior than those who attended schools with less than 1,000 students" (p. 328)
O'Moore et al. (1997)	Primary school	P		n.s.		U	Categories: 0–199, 200–499, >500
	Secondary school	S			+	U	Categories: 0–199, 200–499, >500
Watt (2003)				n.s.			Categories: <400, 400–1000, >1000
Winter (2003)		S			+ + +		Categories: <500, 500–1000, >1000

Table A.19 Overview of studies of school size on attendance/absenteeism and truancy

Overview of studies of school size on attendance/absenteeism and truancy								
Authors	Sample	Country	School type[a]	School size measure	Dependent variable	Mean	SD	Database
Bos et al. (1990)		Netherlands	S	School size	Truancy: the percentage of pupils absent without permission from the total number of potential absentees, during the three days of data collection			Data from 36 schools that participated in the absence registration project
Chen and Weikart (2008)		USA (New York City)	S	Number of students enrolled at each school	Attendance rate: the average percentage of days students come to school for the 2003–2004 school year	960	493	Data from New York City Department of Education (2002–2003 and 2003–2004 school year data for all middle schools)
Durán-Narucki (2008)		USA (New York City, Manhattan borough)	P	The number of students enrolled at each school	Attendance: the average percentage of days that students at a given school attended during the year	712	328	Building Condition Survey (comprehensive study on the condition of New York City school buildings School Report Cards for the year 2000 New York City Board of education
Eccles et al. (1991)		USA (nationally representative)	P and S	Total school enrolment P/K/1–8 th grade schools versus middle and junior high schools (grades 6–8, grades 7–8, grades 7–9) P/K/1–8 th grade schools smaller on the average than other three types	Absenteeism (student and teacher questionnaire): e.g., Student absenteeism is not a problem			National Education Longitudinal Study (NELS: 88)

(continued)

Table A.19 (continued)

Overview of studies of school size on attendance/absenteeism and truancy

Authors	Sample	Country	School type[a]	School size measure	Dependent variable	Mean	SD	Database
Foreman-Peck and Foreman-Peck (2006)		UK (Wales)	S	Log (previous year pupil numbers) School size 1996 School size 2002	% of nonattendance	871 936	331 328	Dataset provided by the school and teacher statistics division of the Welsh assembly government
Gardner et al. (2000)		USA (California)	S	School size Small schools (200–600 pupils) Large schools (>2,000 pupils)	Absenteeism rate for each school	424 2500		Data procured from the education finance division of the California department of education Data were obtained from the 1995/1996 California high school Performance Report
Haller (1992)		USA (nationally representative)	PS	School size: enrolment in 1980	Truancy (reported by principals): a score on the seriousness of two types of attendance problems in their school (unexcused absences class cutting) Truancy (reported by students): % of students who perceived two types of truancy to be a problem in their school Truancy self-(reported: % of students that had personally taken unexcused absences)	963	1219	High School and Beyond surveys (1980 en 1982)

(continued)

Table A.19 (continued)

Overview of studies of school size on attendance/absenteeism and truancy

Authors	Sample	Country	School type[a]	School size measure	Dependent variable	Mean	SD	Database
Heck (1993)		USA (Western state)	PS	Actual size of enrolment	Attendance: percentage of daily attendance			State department of education's survey on restructuring the curriculum
Jones et al. (2008)		USA (Texas)	S	School enrolment	Average daily attendance rate	1012	849	Texas education agency's academic excellence indicator systems
Kahne et al. (2008)		USA (Chicago)	S	School size	Absences: average number of days students were absent from their classes during one academic year			Consortium on Chicago school research's biannual survey. administrative records of CPS and test data
Kuziemko (2006)		USA (Indiana)	P	Abrupt change in school enrolment	Absolute change in average daily attendance rate change over 1 year, 2 years, 3 years	418	170	Indiana department of education: attendance data Public school universe data from national center for educational statistics
Lee et al. (2011)	03–04 04–05 05–06 06–07 07–08	USA (Ohio)	S	Small schools in Ohio versus traditional schools that are identified as similar to the small schools Large school at or above 800 students. Small learning communities approximately 100 students per grade level or 400 students with the learning community	Attendance rate			Data collected from more than 230 Ohio schools

Table A.20 Methodological information available from studies of school size on attendance/absenteeism and truancy

Methodological information available from studies of school size on attendance/absenteeism and truancy											
Authors	Sample	Dependent variable	Number of schools ínclude	Number of classes included	Number of students included	Statistical technique used	Value added	Effects reported	SE reported	Direction of the effect	Further information
Bos et al. (1990)		Truancy				Regression	Yes	n.r.		n.s.	.
Chen and Weikart (2008)		Attendance rate	213			Structural Equation Modelling (school level)	Yes	ß = −0.08		−	
Durán-Narucki (2008)		Attendance	95			Regression (standardized)	Yes	ß = 0.370		+	Mediation model (attendance is mediator)
Eccles et al. (1991)		Absenteeism	759			Regression	No	ß = −0.086		−	"the larger the school size, the more absenteeism … were reported as being a major problem at the school by both teachers and students" (p. 351)
Foreman-Peck and Foreman-Peck (2006)		% of nonattendance Ln school size	1119			Logistic regression (Unstandardized)	Yes	$b = 0.075$		−	No optimum size for attendance
Gardner et al. (2000)		Absenteeism (small vs. large schools)	127			An(c)ova	Yes	F(1.117) = 8.51		−	
Haller (1992)		Truancy Principals Students Self-reported	558			Regression (standardized)	Yes	ß = 0.370 ß = 0.354 ß = 0.335		− − −	

(continued)

Table A.20 (continued)

Methodological information available from studies of school size on attendance/absenteeism and truancy

Authors	Sample	Dependent variable	Number of schools included	Number of classes included	Number of students included	Statistical technique used	Value added	Effects reported	SE reported	Direction of the effect	Further information
Heck (1993)		Attendance	235			Regression (standardized)	Yes	ß = −0.28		−	
Jones (2008)		Attendance rate	1039			Regression (unstandardized)[b] Standardized	Yes	$b = -0.0002$ ß = −0.28		−	
Kahne et al. (2008)		Absences 2002-2003 2003-2004 2004-2005 2005-2006	80			Multilevel (unstandardized)	Yes	Difference −7.4 0.9 1.9 −4.8		− n.s. n.s. − 	
Kuziemko (2006)		Absolute change in average daily attendance rate change over 1 year 2 years 3 years	>100			2 SLS regression (unstandardized)	Yes	$b = 0.003$ $b = -0.003$ $b = -0.004$	0.0015 0.0019 0.0018	n.s. n.s. −	
Lee et al. (2011)	03–04 04–05 05–06 06–07 07–08	Attendance rate (small schools vs. similar schools)	230			Mann-Whitney test	No	– – – – –		n.s. n.s. n.s. n.s. +	"2007–2008 is the only year showing a statistically significant difference in attendance rates between small and similar schools, with similar schools having higher attendance rates" (p. 17)

Table A.21 Overview of studies of school size on drop-out

Overview of studies of school size on drop-out								
Authors	Sample	Country	School type[a]	School size measure	Dependent variable	Mean	SD	Database
Gardner et al. (2000)		USA (California)	S	School size Small schools (200–600 pupils) Large schools (>2000 pupils)	Dropout rate	424 2500		Data procured from the education finance division of the California department of education Data were obtained from the 1995/1996 California high school performance report
Kahne et al. (2008)		USA (Chicago)	S	School size	Drop-out rate: the proportion of students who began as first-time ninth graders at a Chicago Public School, who did not transfer out of the district, who did not graduate, and who were listed as inactive.			Consortium on Chicago school research's biannual survey. Administrative records of CPS and test data

(continued)

Table A.21 (continued)

Overview of studies of school size on drop-out

Authors	Sample	Country	School type[a]	School size measure	Dependent variable	Mean	SD	Database
Lee and Burkam (2003)		USA (nationally representative)	S	Enrolment size Small 0–600 students Medium 601–1,500 students RF Large 1,501–2,500 students Very large >2,500 students	Dropped out between 10th and 12th grade (based on school reports and confirmation form the student or the student's parents			High school effectiveness study (supplementary data collection to NELS: 88)
Rumberger and Palardy (2005)		USA (nationwide)	S	1–600 (small) 601–1,200 (medium) RF 1201–1800 (large) >1,800 (extra large)	Proportion of 10th grade students who dropped out between grades 10–12			National education longitudinal survey (NELS: 88)

Table A.22 Methodological information available from studies of school size on drop-out

Methodological information available from studies of school size on drop-out											
Authors	Sample	Dependent variable	Number of schools ínclude	Number of classes included	Number of students included	Statistical technique used	Value added	Effects reported	SE reported	Direction of the effect	Further information
Gardner et al. (2000)		Dropout rate (small vs. large schools)	127			An(c)ova	Yes	F(1.117) = 7.25		–	
Kahne et al. (2008)		Dropout rate 2002–2003 2003–2004	80			Multilevel (unstandardized)	Yes			n.s. n.s.	
Lee and Burkam (2003)		Dropped out between 10th and 12th grade 0–600 (small) 601–1,500 (medium) RF 1,501–2,500 (large) >2,500 (very large) 0–600 (small) 601–1,500 (medium) RF 1,501–2,500 (large) >2,500 (very large)	190		2,480	Multilevel (logistic) (unstandardized)	Yes	Change in log odds[b] 0.75 1.32 0.76 Change in odds 2.12 3.74 2.14			Positive odds associated with greater likelihood of dropping out. Compared to medium sized schools, large and very large schools had significantly higher drop-out rates. The highest drop-out rate was found in large schools. Small school also had higher drop-out rates (significant at 10 % level)

(continued)

Table A.22 (continued)

Methodological information available from studies of school size on drop-out

Authors	Sample	Dependent variable	Number of schools included	Number of classes included	Number of students included	Statistical technique used	Value added	Effects reported	SE reported	Direction of the effect	Further information
Rumberger and Palardy (2005)		Drop-out 1–600 (small) 601–1,200 RF 1,201–1,800 (large) >1,800 (extra large)	912		14,199	Two level multinomial logistic regression	Yes	School effect size −0.54 0.227 0.145		∩ 1200–1800	School effect size computed by first converting HLM coefficients to standard units and then dividing by the school-level standard deviation of the dependent variable estimated from the HLM null model

The results are presented in the log odds metric. Since this metric is not easily interpretable, the results were interpreted into an odds ratio (the ratio between p, the probability of dropping out, and $1 - p$, the probability of remaining in school). The odds ratio permits an estimate of the percentage increase or decrease in the odds of dropping out. For example, a change in the odds ratio of 1.75 represents a 75 % increase in the likelihood (or odds) of dropping out. A change in the odds ratio of 0.40 represents a 60 % decrease in the likelihood of dropping out (p. 373)

Table A.23 Results of vote counts examining the number of negative, nonsignificant, and positive effects of school size on attendance/absenteeism and truancy for each sample (school size measured as a continuous variable)

Study	Sample	Education level	Negative	Not significant	Positive	Total
Bos et al. (1990)		S	0	1	0	1
Chen and Weikart (2008)		S	0	1	0	1
Durán-Narucki (2008)		P	0	0	1	1
Eccles et al. (1991)		PS	1	0	0	1
Foreman-Peck and Foreman-Peck (2006)		S	1	0	0	1
Haller (1992)		S	3	0	0	3
Heck (1993)		PS	0	1	0	1
Jones et al. (2008)		S	1	0	0	1
Kahne et al. (2008)		S	2	2	0	4
Kuziemko (2006)		S	0	2	1	3
Lee et al. (2011)	2003–2004	S	0	1	0	1
	2004–2005	S	0	1	0	1
	2005–2006	S	0	1	0	1
	2006–2007	S	0	1	0	1
	2007–2008	S	1	0	0	1
Total			9	11	2	22

Table A.24 Results of vote counts examining the number of negative, nonsignificant, positive, and curvilinear effects of school size on attendance/absenteeism and truancy for each sample (school size measured as categories)

Study	Sample	Education level	Direction of effect				Remarks
			−	n.s.	+	Curvilinear	
Gardner et al. (2000)		S	−				Small schools (200–600) versus large schools (>2,000)

Table A.25 Results of vote counts examining the number of negative, nonsignificant, and positive effects of school size on drop out for each sample sample (school size measured as a continuous variable)

Study	Sample	Education level	Negative	Not significant	Positive	Total
Kahne et al. (2008)	2002–2003	S	0	1	0	1
	2003–2004		0	1	0	1
Total			0	2	0	2

Table A.26 Results of vote counts examining the number of negative, nonsignificant, positive, and curvilinear effects of school size on drop-out for each sample (school size measured as categories)

Study	Sample	Education level	Direction of effect				Remarks
			−	n.s.	+	Curvilinear	
Gardner et al. (2000)		S	−				Small schools (200–600) versus large schools (>2,000)
Lee and Burkam (2003)		S				U 601–1500	Categories: 0–600, 601–1500 (RF), 1501–2500, >2500 Compared to medium sized schools, large and very large schools had significantly higher drop-out rates. The highest drop-out rate was found in large schools. Small school also had higher drop-out rates (significant at 10 % level)
Rumberger and Palardy (2005)		S				∩ 1200–1800	Categories: 0–600, 601–1200 (RF), 1201–1800, >180

Table A.27 Overview of studies of school size on other student outcomes

Overview of studies of school size on other student outcomes								
Authors	Sample	Country	School type[a]	School size measure	Dependent variable	Mean	SD	Database
Coladarci and Cobb (1996)		USA (nationally representative)	S	School size Compares smaller (<800) and larger (>= 1,600) schools Students from schools with other sizes eliminated from analyses	Self-esteem (mean across seven items administered in senior year (e.g., I feel I am a person of worth)			National education longitudinal study of 1988 (NELS: 88)
Holas and Huston (2012)	Grade 6	USA (nationally)	P	Total enrolment	Perceived self-competence: scale 12 items focussing on perception of efficacy and competence in English and math	690	300	NICHD study of early child care and youth development
Kirkpatrick Johnson et al. (2001)	Middle schools High schools	USA (nationally representative)	S	Total enrolment in schools in hundreds of students	Engagement in school: three items (reversed coded): the past school year how many times the adolescents had skipped school, had trouble paying attention in school and had trouble getting homework done	477 1147	234 716	National longitudinal study of adolescent health

(continued)

Table A.27 (continued)

Overview of studies of school size on other student outcomes

Authors	Sample	Country	School type[a]	School size measure	Dependent variable	Mean	SD	Database
Lay (2007)		USA (nationwide)	S	School size Continuous measure Categories based on continuous measure <300, 301–600, 601–900, 901–1200, 1201–1500, 1501–1800, >1800 Categories based on parental responses <300, 300–599, 600–999, >1000	Participation in community services			National household education survey
Lee and Smith (1995)		USA (nationally representative)	S	Total enrolment of as October 1989 (transformed to its natural logarithm and standardized)	10th grade academic Engagement: standardized factor weighted composite of eight items measuring students' behaviors (e.g., often work hard in math class, often feel challenged in math class)			National Education Longitudinal Study of 1988 (NELS: 88) 1st and 2nd wave Mean school size: Traditional schools: 1095 Moderate schools: 633 Restructuring schools: 764

(continued)

Table A.27 (continued)

Overview of studies of school size on other student outcomes

Authors	Sample	Country	School type[a]	School size measure	Dependent variable	Mean	SD	Database
Weiss et al. (2010)		USA (nationally representative)	S	Total number of students in school Categories: Small: 1–599 students (RF) Moderately small 600–999 Moderately large 1,000–1,599 Large 1,600–2,499	School engagement: composite of students 'psychological and behavioral connections with the values and aims of school (seven variables included: teacher experience, delinquent behavior, academic friend, educational motivation, teachers' belief about ability, school preparedness, parental involvement)			Educational longitudinal study of 2002 (ELS:2002) 10th grade

Table A.28 Methodological information available from studies of school size on other student outcomes

Methodological information available from studies of school size on other student outcomes

Authors	Sample	Dependent variable	Number of schools ínclud	Number of classes included	Number of students included	Statistical technique used	Value added	Effects reported	SE reported	Direction of the effect	Further information
Coladarci and Cobb (1996)		Self-esteem			4567	Structural equation modelling (individual level)	Yes	ß = −0.015		n.s.	Compares students from smaller (<800) and larger (>= 1,600) schools
Holas and Huston (2012)	Grade 6	Perceived self-competence			828	Structural equation modelling	Yes	n.r.		n.s.	.
Kirkpatrick Johnson et al. (2001)	Middle schools High schools	Engagement in school	45 64		2,482 8,104	Multilevel (standardized)	Yes	n.r. ß = −0.07		n.s. −	
Lay (2007)		Participation in community services School size as continuous measure Categories based on continuous measure <300 301–600 601–900 901–1,200 1,501–1,800 >1,800 Categories based on parental responses <300 600–999 >1,000			3,010	Logistic regression	Yes	$b = 0.00001$ $b = 0.427$ $b = 0.001$ $b = 0.265$ $b = -0.031$ $b = -0.033$ $b = 0.114$ $b = 0.357$ $b = 0.032$ $b = 0.070$	0.00004 0.165 0.140 0.138 0.135 0.134 0.115 0.163 0.107 0.091	n.s.	Students in schools with fewer than 300 students significantly more likely volunteering in community service Students in schools with fewer than 300 students significantly more likely volunteering in community service

(continued)

Table A.28 (continued)

Methodological information available from studies of school size on other student outcomes

Authors	Sample	Dependent variable	Number of schools íncluded	Number of classes included	Number of students included	Statistical technique used	Value added	Effects reported	SE reported	Direction of the effect	Further information
Lee and Smith (1995)		Academic engagement	820		11,794	Multilevel (standardized)	Yes	ES = −0.19		−	
Weiss et al. (2010)		School engagement Small 1–599 (RF) Moderately small 600–999 Moderately large 1,000–1,599 Large 1,600–2,499			10,946	Multilevel					(unstandardized)
Yes	$b = -0.065$ $b = -0.135$ $b = -0.146$	0.045 0.044 0.030	−								

Table A.29 Results of vote counts examining the number of negative, nonsignificant, and positive effects of school size on other student outcome variables for each sample (school size measured as a continuous variable)

Study	Dependent variable	Sample	Education level	Negative	Not significant	Positive	Total
Holas and Huston (2012)	Perceived self-competence		P	0	1	0	1
Kirkpatrick Johnson et al. (2001)	Engagement in schools	Middle schools	S	0	1	0	1
		High schools		1	0	0	1
Lay (2007)	Participation in community services		S	0	1	0	1
Lee and Smith (1995)	Academic engagement		S	1	0	0	1
Total				2	3	0	5

Table A.30 Results of vote counts examining the number of negative, nonsignificant, positive, and curvilinear effects of school size on other student outcomes for each sample (school size measured as categories)

Study	Dependent variable	Education level	Direction of effect				Remarks
			−	n.s.	+	Curvilinear	
Coladarci and Cobb (1996)	Self-esteem	S		n.s.			Categories: <800, >1599
Lay (2007)	Participation in community services	S				Students in schools with fewer than 300 students significantly more likely volunteering in community service	Categories based on continuous measure: <300, 301–600, 601–900, 901–1200, 1501–1800, >1800
						Students in schools with fewer than 300 students significantly more likely volunteering in community service	Categories based on parental responses: <300, 600–999, >1000
Weiss et al. (2010)	School engagement	S	−				Categories: 1–599, 600–999, 1000–1599, 1600–2499

Table A.31 Overview of studies of school size on school organisation and teaching and learning

Overview of studies of school size on school organisation and teaching and learning								
Authors	Sample	Country	School type[a]	School size measure	Dependent variable	Mean	SD	Database
Eccles et al. (1991)		USA (nationally representative)	P and S	Total school enrolment P/K/1–8th grade schools versus middle and junior High schools (grades 6–8, grades 7–8, grades 7–9) P/K/1–8th grade schools smaller on the average than other three types	Teacher efficacy			National education longitudinal study (NELS: 88)
Inspectorate of Education (2003)		Netherlands	S	School size <500 501–1,000 >1,000	Teaching-learning process: pedagogic and didactic approach			Data from regular supervision of schools (years 1999–2000, 2000–2001)
Kahne et al. (2008)		USA (Chicago)	S	School size	Collective responsibility: teachers' assessment of the strength of their shared commitment to improve the school so that all students learn. Commitment to innovation Expectations for postsecondary education: teachers' reports of the degree to which they expect that most students at their school will go to college			Consortium on Chicago School Research's biannual survey. Administrative records of CPS and test data

(continued)

Table A.31 (continued)

Overview of studies of school size on school organisation and teaching and learning

Authors	Sample	Country	School type[a]	School size measure	Dependent variable	Mean	SD	Database
					Principal instructional leadership: teachers' perception of their principals as instructional leaders			
					Program coherence: the degree to which teachers believe that the programs at their schools are coordinated with one another and with the school's mission			
					Quality professional development			
					Quality student discussions in classroom			
					Reflective dialogue: teachers' assessment of how often teachers talk with one another about instruction and student learning			
					Academic press: students' views of their teachers' efforts to push students to higher levels of academic performance			

(continued)

Table A.31 (continued)

Overview of studies of school size on school organisation and teaching and learning

Authors	Sample	Country	School type[a]	School size measure	Dependent variable	Mean	SD	Database
					Quality English instruction: student reports of the frequency with which students are made to practice higher order english activities			
					Quality math instruction: student reports of the frequency with which students are made to practice higher-order math activities			
					Peer support for academic achievement: the norms among students with regard to their peers' support of academic work			
					School-wide future orientation: student reports of the degree to which (a) teachers work hard to make sure that all students are learning, are staying in school, are planning for their futures and (b) all students are encouraged to go to college			

(continued)

Table A.31 (continued)

Overview of studies of school size on school organisation and teaching and learning

Authors	Sample	Country	School type[a]	School size measure	Dependent variable	Mean	SD	Database
Silins and Mulford (2004)		Australia	S	School size in 1997	Organisational Learning—the extent to which the school is perceived to be functioning as a learning organisation according to measures on the four factors that define organisational learning: collaborative climate, Taking initiatives and risks, Improving school practices, Professional development Teacher Leadership—the extent to which individual teachers, teacher teams or committees and whole staff working together are a source of leadership in the school Teachers' work—the construct representing students' perceptions of teachers' work in the classroom including their liking of the way teachers instruct, the variety of instructional activities employed, the extent teachers discuss students' work with them, the organization of their classes, the expectations that they will do their best work, and the extent students are challenged in class	632	283	Leadership for organizational learning and student outcomes (LOLSO)

Table A.32 Methodological information available from studies of school size on school organisation and teaching and learning

Methodological information available from studies of school size on school organisation and teaching and learning											
Authors	Sample	Dependent variable	Number of schools ínclude	Number of classes included	Number of students included	Statistical technique used	Value added	Effects reported	SE reported	Direction of the effect	Further information
Eccles et al. (1991)		Teacher efficacy	759			Regression	No	ß = −0.032		−	"the larger the school size, the less efficacious … and … were reported as being a major problem at the school by both teachers and students" (p. 351)
Inspectorate of Education (2003)		Pedagogic and didactic approach	378			ANOVA	No			500–1,000 students lowest score	
Kahne et al. (2008)		Collective responsibility (teacher)	80			Multilevel		(unstandardized)	Yes	$b = 0.486$	
−		Commitment to innovation (teacher)						$b = -0.529$		−	
		Expectations for postsecondary education (teacher)						$b = -0.938$		−	
		Principal instructional leadership (teacher)						$b = 0.272$		n.s.	−
		Program coherence (teacher)						$b = -0.184$		n.s.	

(continued)

Table A.32 (continued)

Methodological information available from studies of school size on school organisation and teaching and learning

Authors	Sample	Dependent variable	Number of schools ínclud	Number of classes included	Number of students included	Statistical technique used	Value added	Effects reported	SE reported	Direction of the effect	Further information
		Quality professional development (teacher)						$b = 0.038$		n.s.	–
		Quality student discussion (teacher)						$b = -0.108$		n.s.	–
		Reflective dialogue (teacher)						$b = -0.081$		n.s.	
		Academic press (student)						$b = -0.187$		n.s.	–
		Quality English instruction (student)						$b = -0.036$		n.s.	
		Quality math instruction (student)						$b = -0.009$		n.s.	
		Peer support for academic achievement (student)						$b = -0.559$		–	
		School-wide future orientation (student)						$b = -0.326$		n.s.	
Silins and Mulford (2004)		Teacher leadership	96	2503	3500	Structural Equation Modelling	Yes	ß = −0.15		n.s.	
		Organisational learning						ß = −0.23		–	
		Teachers' work						ß = −0.06		n.s.	

Table A.33 Results of vote counts examining the number of negative, nonsignificant, and positive effects of school size on school organization and teaching and learning for each sample (school size measured as a continuous variable)

Study	Dependent variable	Sample	Education level	Negative	Not significant	Positive	Total
Eccles et al. (1991)	Teacher efficacy		PS	1	0	0	1
Kahne et al. (2008)	Collective responsibility		S	1	0	0	1
	Commitment to innovation			1	0	0	1
	Expectations for postsecondary education			1	0	0	1
	Principal instructional leadership			0	1	0	1
	Program coherence			0	1	0	1
	Quality professional development			0	1	0	1
	Quality student discussions in classroom			0	1	0	1
	Reflective dialogue			0	1	0	1
	Academic press			0	1	0	1
	Quality English instruction			0	1	0	1
	Quality Math instruction			0	1	0	1
	Peer support for academic achievement			1	0	0	1
	School-wide future orientation			0	1	0	1
Silins and Mulford (2004)	Organisational learning		S	1	0	0	1
	Teacher leadership			0	1	0	1
	Teachers' work in the classroom			0	1	0	1
Total				6	11	0	17

Table A.34 Results of vote counts examining the number of negative, nonsignificant, positive, and curvilinear effects of school size on teaching and learning for each sample (school size measured as categories)

Study	Dependent variable	Education level	Direction of effect				Remarks
			−	n.s.	+	Curvilinear	
Inspectorate of Education (2003)	Teaching-learning process	S				500-1000 students lowest score	Categories: <500, 501–1,000 >1,000
Lay (2007)	Participation in community services	S				Students in schools with fewer than 300 students significantly more likely volunteering in community service	Categories based on continuous measure: <300, 301–600, 601–900, 901–1200 1501–1800, >1800
						Students in schools with fewer than 300 students significantly more likely volunteering in community service	Categories based on parental responses: <300, 600–999 >1,000
Rumberger and Palardy (2005)	Transfer	S				U 1200–1800	Categories: 0–600, 601–1200 (RF), 1201–1800, >180

Table A.35 Overview of studies of school size costs

Overview of studies of school size costs

Authors	Sample	Country	School type[a]	School size measure	Costs or cost-efficiency measure	Mean	SD	Database
Bickel et al. (2001)		USA	S	Number of students: natural logarithms of single-student units	Expenditure per pupil	877	850	Texas dataset of 1,001 high schools
Bowles and Bosworth (2002)		USA	PS	Average daily membership for school *i* for period *t* (Natural logarithm)	Operating expenditures per student in school *i* for period *t* (Natural logarithm)			Data from 17 Wyoming school districts
Lewis and Chakraborty (1996)		USA	PS	Number of students per school, average 1982–1993 (Natural logarithm)	Operating expenditure per student 1982–1993 (Natural logarithm)	511 (median)		Data from 40 Utah school districts
Merkies (2000)		Netherlands	P	Number of pupils (Natural logarithm)	Total costs of a school/costs per pupil (Natural logarithm)	200		Dataset compromising 1784 primary schools in the Netherlands in the year 1986/1987
Stiefel et al. (2000)		USA	P	School size: natural logarithm of number of 1995–1996 registered general education students	Natural logarithm of budget per graduate: 1995-1996 total budget per student, multiplied by 4, adjusted. Graduate: from cohort of 9th graders, number who graduated from school in 4 years (transfers in attributed to last school attended, transfers out of New York system removed form cohort)	2030	1192	Data from board of education from the city of New York

Table A.36 Methodological information available from studies of school size on costs

Methodological information available from studies of school size on costs											
Authors	Sample	Costs measure	Number of schools ínclude	Number of classes included	Number of students included	Statistical technique used	Value added	Effect(s) reported	SE reported	Direction of the effect	Further information
Bickel et al. (2001)		Expenditure per student Size expressed in ln of single unit student	1,001			Regression (standardized)	Yes	ß = −0.199		–	Grade span configuration included in the analysis: K-12 "unit schools" covering all grade levels more cost effective than traditional high schools
Bowles and Bosworth (2002)		Operating expenditures per student	80			Regression (simultaneous equation modelling) unstandardized	Yes	$b = -0.2052$		–	
Lewis and Chakraborty (1996)		Operating expenditure per student				Regression	Yes	$b = -0.15508$			Both school size and district size (together with covariates) included in the analysis. Only the school size effect is found to be significant
Merkies (2000)		Average school costs	1,784			Regression	Yes				"…. considerable economies of scale can be acquired by small schools. These benefits dissipate as schools get larger. From the average school (200 pupils) onwards the average costs remain virtually constant. For schools with more than twice the average number of pupils there are no more economies of scale. The optimal size is around 450 pupils" (p. 206)
Stiefel et al. (2000)		Budget per graduate	121			Regression (unstandardized)	Yes	$b = -0.140$	0.048	–	

Table A.37 Results of vote counts examining the number of negative, nonsignificant, and positive effects of school size on costs for each sample (school size measured as a continuous variable)

Study	Sample	Education level	Negative	Not significant	Positive	Total
Bickel et al. (2001)		S	1	0	0	1
Bowles and Bosworth (2002)		PS	1	0	0	1
Lewis and Chakraborty (1996)		PS	1	0	0	1
Merkies (2000)		P	1	0	0	1
Stiefel et al. (2000)		P	0	1	0	1
Total			4	1	0	5

Note all relations relations are modeled as log-linear functions. An adequate interpretation of this is given by Merkies (2000, p. 206): "… considerable economies of scale can be acquired by small schools. These benefits dissipate as schools get larger"

References

Andrews, M., Duncombe, W., & Yinger, J. (2002). Revisiting economies of size in American education: Are we any closer to a consensus? *Economics of Education Review,21*(3), 245–262.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: Wiley.

Bryk, A. S., & Thum, Y. M. (1989). The effects of high school organization on dropping out: An exploratory investigation. *American Educational Research Journal,26*(3), 353–383.

Bushman, B. J. (1994). Vote-counting procedures in meta-analysis. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis*. New York: Russell Sage Foundation.

Bushman, B. J., & Wang, M. C. (2009). Vote-counting procedures in meta-analysis. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 208–222). New York: Russell Sage Foundation.

Chakraborty, K., Biswas, B., & Lewis, W. C. (2000). Economies of scale in public education: An econometric analysis. *Contemporary Economic Policy,18*(2), 238–247.

Cooper, H., Hedges, L. V., & Valentine, J. C. (Eds.). (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). New York: Russell Sage Foundation.

Gutman, L. M., & Midgley, C. (2000). The role of protective factors in supporting the academic achievement of poor African-American students during the middle school transition. *Journal of Youth and Adolescence, 29*(2), 223–248. doi:10.1023/A:1005108700243

Haller, E. J., Monk, D. H., Spotted Bear, A., Griffith, J. & Moss, P. (1990). School size and program cohesiveness: Evidence from High School and Beyond. *Educational Evaluation and Policy Analysis, 12,* 109–120.

Hendriks, M., Scheerens, J., & Steen, R. (2008). *Schaalgrootte en de menselijke maat*. Enschede: Universiteit Twente.

Leithwood, K., & Jantzi, D. (2009). A review of empirical evidence about school size effects: A policy perspective. *Review of Educational Research,79*(1), 464–490. doi:10.3102/0034654308326158

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Monk, D. H. (1994). Secondary school size and curriculum comprehensiveness. *Economics of Education Review, 6*, 137–150.

Newman, M., Garrett, Z., Elbourne, D., Bradley, S., Noden, P., Taylor, J., et al. (2006). Does secondary school size make a difference?. A systematic review. *Educational Research Review,1*(1), 41–60. doi:10.1016/j.edurev.2006.03.001

Newmann, F., Wehlage, G., & Lamborn, S. (1992) The significance and sources of student engagement. In Newmann (Eds.), *Student engagement and achievement in American secondary schools* (p. 11–39). New York: Teachers College Press.

Rumberger, R. W. (1995). Dropping out of middle school: A multilevel analysis of students and schools. *American Educational Research Journal, 32*(3), 583–625.

Scheerens, J., Seidel, T., Witziers, B., Hendriks, M., & Doornekamp, G. (2005). *Positioning and validating the supervision framework*. Enschede: University of Twente, Department of Educational Organization and Management.

Scheerens, J., Luyten, H., Steen, R., & Luyten-de Thouars, Y. (2007). *Review and meta analyses of school and teaching effectiveness*. Enschede: Department of Educational Organisation and Management, University of Twente.

Spielhofer, T., Benton, T., & Schagen, S. (2004). A study of the effects of school size and single-sex education in English schools. *Research Papers in Education,19*(2), 133–159. doi:10.1080/02671520410001695407

Welsh, W. N., Stokes, R., & Greene J. R. (2000). A macro-level model of school disorder. *Journal of Research in Crime and Delinquency, 37*(3), 243–283.

Studies Used for Vote Count

Åberg-Bengtsson, L. (2004). Do small rural schools differ? A comparative two-level model of reading achievement among Swedish 9-year-olds. *Scandinavian Journal of Educational Research,48*(1), 19–33. doi:10.1080/0031383032000149823

Alspaugh, J. W. (2004). School size as a factor in elementary school achievement. *ERS Spectrum,22*(2), 28–34.

Archibald, S. (2006). Narrowing in on educational resources that do affect student achievement. *Peabody Journal of Education, 81*(4), 23–42. http://dx.doi.org/10.1207/s15327930pje8104_2

Attar-Schwartz, S. (2009). Peer sexual harassment victimization at school: The roles of student characteristics, cultural affiliation, and school factors. *American Journal of Orthopsychiatry,79*(3), 407–420. doi:10.1037/a0016553

Barnes, J., Belsky, J., Broomfield, K. A., & Melhuish, E. (2006). Neighbourhood deprivation, school disorder and academic achievement in primary schools in deprived communities in England. *International Journal of Behavioral Development,30*(2), 127–136. doi:10.1177/0165025406065385

Bickel, R., Howley, C., Williams, T., & Glascock, C. (2001). High school size, achievement equity, and cost: Robust interaction effects and tentative results. *Education Policy Analysis Archives, 9*(40). Retrieved October 17, 2012 from http://epaa.asu.edu/ojs/article/view/369/495

Bonnet, M., Gooss, F. A., Willemen, A. M., & Schuengel, C. (2009). Peer victimization in dutch school classes of four- to five-year-olds: Contributing factors at the school level. *The Elementary School Journal,110*(2), 163–177.

Borland, M. V., & Howsen, R. M. (2003). An examination of the effect of elementary school size on student academic achievement. *International Review of Education,49*(5), 463–474.

Bos, K. T., Ruijters, A., & Visscher, A. (1990). Truancy, drop-out, class repeating and their relation with school characteristics. *Educational Research,32*(3), 175–185.

Bowen, G. L., Bowen, N. K., & Richman, J. M. (2000). School size and middle school students' perceptions of the school environment. *Social Work in Education,22*(2), 69–82.

Bowes, L., Arseneault, L., Maughan, B., Taylor, A., Caspi, A., & Moffitt, T. E. (2009). School, neighborhood, and family factors are associated with children's bullying involvement: A nationally representative longitudinal study. *Journal of the American Academy of Child and Adolescent Psychiatry,48*(5), 545.

Bowles, T. J., & Bosworth, R. (2002). Scale economies in public education: Evidence from school level data. *Journal of Education Finance,28*(2), 285–299.

Bradley, S., & Taylor, J. (1998). The effect of school size on exam performance in secondary schools. *Oxford Bulletin of Economics and Statistics,60*(3), 291–324. doi:10.1111/1468-0084.00102

Caldas, S. J. (1993). Reexamination of input and process factor effects on public school achievement. *The Journal of Educational Research,86*(4), 206–214. doi:10.1080/00220671.1993.9941832

Carolan, B. V. (2012). An examination of the relationship among high school size, social capital, and adolescents' mathematics achievement. *Journal of Research on Adolescence,22*(3), 583–595. doi:10.1111/j.1532-7795.2012.00779.x

Chen, G. (2008). Communities, students, schools and school crime—A confirmatory study of crime in US high schools. *Urban Education,43*(3), 301–318. doi:10.1177/0042085907311791

Chen, P., & Vazsonyi, A. T. (2013). Future orientation, school contexts, and problem behaviors: A multilevel study. *Journal of Youth and Adolescence,42*(1), 67–81. doi:10.1007/s10964-012-9785-4

Chen, G., & Weikart, L. A. (2008). Student background, school climate, school disorder, and student achievement: An empirical study of New York city's middle schools. *Journal of School Violence, 7*(4), 3–20. http://dx.doi.org/10.1080/15388220801973813

Coladarci, T., & Cobb, C. D. (1996). Extracurricular participation, school size, and achievement and self-esteem among high school students: A national look. *Journal of Research in Rural Education,12*(2), 92–103.

Crosnoe, R., Kirkpatrick Johnson, M., & Elder, G. H. (2004). School size and the interpersonal side of education: An examination of race/ethnicity and organizational context. *Social Science Quarterly*, *85*(5), 1259–1274. doi: 10.1111/j.0038-4941.2004.00275.x

Dee, T. S., Ha, W., & Jacob, B. A. (2007). *The effects of school size on parental involvement and social capital: Evidence from the ELS:2002*. Brookings papers on Education Policy, pp 77–97.

Deller, S. C., & Rudnicki, E. (1993). Production efficiency in elementary education: The case of Maine public schools. *Economics of Education Review,12*(1), 45–57.

Driscoll, D., Halcoussis, D., & Svorny, S. (2003). School district size and student performance. *Economics of Education Review,22*(2), 193–201. doi:10.1016/S0272-7757(02)00002-X

Durán-Narucki, V. (2008). School building condition, school attendance, and academic achievement in New York City public schools: A mediation model. *Journal of Environmental Psychology,28*(3), 278–286. doi:10.1016/j.jenvp.2008.02.008

Eberts, R. W., Schwartz, E. K., & Stone, J. A. (1990). School reform, school size, and student achievement. *Economic Review,26*(2), 2–15.

Eccles, J. S., Lord, S., & Midgely, C. (1991). What are we doing to early adolescents? The impact of educational contexts on early adolescents. *American Journal of Education,99*(4), 521–542.

Feldman, A. F., & Matjasko, J. L. (2007). Profiles and portfolios of adolescent school-based extracurricular activity participation. *Journal of Adolescence,30*(2), 313–332. doi:10.1016/j.adolescence.2006.03.004

Fernandez, K. E. (2011). Evaluating school improvement plans and their affect on academic performance. *Educational Policy,25*(2), 338–367. doi:10.1177/0895904809351693

Foreman-Peck, J., & Foreman-Peck, L. (2006). Should schools be smaller? The size-performance relationship for Welsh schools. *Economics of Education Review,25*(2), 157–171. doi:10.1016/j.econedurev.2005.01.004

Fowler, W. J., & Walberg, H. J. (1991). School size, characteristics, and outcomes. *Educational Evaluation and Policy Analysis,13*(2), 189–202. doi:10.2307/1164583

Gardner, P. W., Ritblatt, S. N., & Beatty, J. R. (2000). Academic achievement and parental involvement as a function of high school size. *The High School Journal,83*(2), 21–27.

Gottfredson, D. C., & DiPietro, S. M. (2011). School size, social capital, and student victimization. *Sociology of Education,84*(1), 69–89. doi:10.1177/0038040710392718

Haller, E. J. (1992). High-school size and student indiscipline: Another aspect of the school consolidation issue. *Educational Evaluation and Policy Analysis,14*(2), 145–156.

Heck, R. H. (1993). School characteristics, school academic indicators and student outcomes: Implications for policies to improve schools. *Journal of Education Policy,8*(2), 143–154.
Holas, I., & Huston, A. C. (2012). Are middle schools harmful? The role of transition timing, classroom quality and school characteristics. *Journal of Youth and Adolescence,41*(3), 333–345. doi:10.1007/s10964-011-9732-9
Inspectie van het Onderwijs. (2003). *Schoolgrootte en kwaliteit. Groot in kleinschaligheid*. Utrecht: Inspectie van het Onderwijs.
Jones, J. T., Toma, E. F., & Zimmer, R. W. (2008). School attendance and district and school size. *Economics of Education Review,27*(2), 140–148. doi:10.1016/j.econedurev.2006.09.005
Kahne, J. E., Sporte, S. E., De La Torre, M., & Easton, J. Q. (2008). Small high schools on a larger scale: The impact of school conversions in Chicago. *Educational Evaluation and Policy Analysis,30*(3), 281–315. doi:10.3102/0162373708319184
Khoury-Kassabri, M., Benbenishty, R., Astor, R. A., & Zeira, A. (2004). The contributions of community, family, and school variables to student victimization. *American Journal of Community Psychology,34*(3–4), 187–204.
Kirkpatrick Johnson, M. K., Crosnoe, R., & Elder, G. H, Jr. (2001). Students' attachment and academic engagement: The role of race and ethnicity. *Sociology of Education, 74*(4), 318–340.
Klein, J., & Cornell, D. (2010). Is the link between large high schools and student victimization an illusion? *Journal of Educational Psychology,102*(4), 933–946. doi:10.1037/a0019896
Koth, C. W., Bradshaw, C. P., & Leaf, P. J. (2008). A multilevel study of predictors of student perceptions of school climate: The effect of classroom-level factors. *Journal of Educational Psychology,100*(1), 96–104. doi:10.1037/0022-0663.100.1.96
Kuziemko, I. (2006). Using shocks to school enrollment to estimate the effect of school size on student achievement. *Economics of Education Review,25*(1), 63–75. doi:10.1016/j.econedurev.2004.10.003
Lamdin, D. J. (1995). Testing for the effect of school size on student achievement within a school district. *Education Economics,3*(1), 33–42.
Lay, J. C. (2007). Smaller isn't always better: School size and school participation among young people. *Social Science Quarterly,88*(3), 790–815.
Lee, V. E., & Burkam, D. T. (2003). Dropping out of high school: The role of school organization and structure. *American Educational Research Journal,40*(2), 353–393.
Lee, V. E., & Loeb, S. (2000). School size in Chicago elementary schools: Effects on teachers' attitudes and students' achievement. *American Educational Research Journal,37*(1), 3–31.
Lee, H. J., Özgün-Koca, S. A., & Cristol, D. (2011). An analysis of high school transformation effort from an outcome perspective. *Current Issues in Education, 14*(1), 1–33. Retrieved on October 12, 2012, from http://cie.asu.edu/ojs/index.php/cieatasu/article/view/
Lee, V. E., & Smith, J. B. (1995). Effects of high school restructuring and size on early gains in achievement and engagement. *Sociology of Education,68*(4), 241–270. doi:10.2307/2112741
Lee, V. E., & Smith, J. B. (1997). High school size: Which works best and for whom? *Educational Evaluation and Policy Analysis,19*(3), 205–227.
Leung, A., & Ferris, J. S. (2008). School size and youth violence. *Journal of Economic Behavior and Organization,65*(2), 318–333. doi:10.1016/j.jebo.2005.10.001
Lewis, W. C., & Chakraborty, K. (1996). Scale economics in public education. *Regional Analysis and Policy,26*(1), 23–35.
Lubienski, S. T., Lubienski, C., & Crane, C. C. (2008). Achievement differences and school type: The role of school climate, teacher certification, and instruction. *American Journal of Education,115*(1), 97–138.
Luyten, H. (1994). School size effects on achievement in secondary education: Evidence from the Netherlands, Sweden, and the USA. *School Effectiveness and School Improvement,5*(1), 75–99. doi:10.1080/0924345940050105
Ma, X., & McIntyre, L. J. (2005). Exploring differential effects of mathematics courses on mathematics achievement. *Canadian Journal of Education/Revue Canadienne de l'éducation,28*(4), 827–852.

Maerten-Rivera, J., Myers, N., Lee, O., & Penfield, R. (2010). Student and school predictors of high-stakes assessment in science. *Science Education,94*(6), 937–962. doi:10.1002/sce.20408

McMillen, B. J. (2004). School size, achievement, and achievement gaps. *Education Policy Analysis Archives, 12*, 1–26. Retrieved on October 12, 2012, from http://epaa.asu.edu/epaa/v12n58/

MacNeal, R. B, Jr. (2008). Participating in high school extracurricular activities: Investigating school effects. *Social Science Quarterly, 80*(2), 291–309.

McNeely, C. A., Nonnemaker, J. M., & Blum, R. W. (2002). Promoting school connectedness: Evidence from the national longitudinal study of adolescent health. *Journal of School Health,72*(4), 138–146.

Merkies, A. H. Q. M. (2000). Economics of scale and school consolidation in dutch primary school industry. In J. L. T. Blank (Ed.), *Public provision and performance: Contributions from efficiency and productivity measurement* (pp. 191–218). Amsterdam, New York and Oxford: Elsevier Science, North-Holland.

Moe, T. M. (2009). Collective bargaining and the performance of the public schools. *American Journal of Political Science,53*(1), 156–174.

Mooij, T., Smeets, E., & de Wit, W. (2011). Multi-level aspects of social cohesion of secondary schools and pupils' feelings of safety. *British Journal of Educational Psychology,81*(3), 369–390. doi:10.1348/000709910X526614

O'Moore, A. M., Kirkham, C., & Smith, M. (1997). Bullying behaviour in Irish schools: A nationwide study. *Irish Journal of Psychology,18*(2), 141–169. doi:10.1080/03033910.1997.10558137

Payne, A. A. (2012). Communal school organization effects on school disorder: Interactions with school structure. *Deviant Behavior,33*(7), 507–524. doi:10.1080/01639625.2011.636686

Ready, D. D., & Lee, V. E. (2006). *Optimal context size in elementary schools: disentangling the effects of class size and school size and school size*. Brookings papers on Education Policy, pp. 99–135.

Rosenblatt, Z. (2001). Teachers' multiple roles and skill flexibility: Effects on work attitudes. *Educational Administration Quarterly,37*(5), 684–708. doi:10.1177/00131610121969479

Rumberger, R. W., & Palardy, G. J. (2005). Test scores, dropout rates, and transfer rates as alternative indicators of high school performance. *American Educational Research Journal,42*(1), 3–42.

Sandy, J., & Duncan, K. (2010). Examining the achievement test score gap between urban and suburban students. *Education Economics,18*(3), 297–315. doi:10.1080/09645290903465713

Sawkins, J. W. (2002). Examination performance in Scottish secondary schools: An ordered logic approach. *Applied Economics, 34*(16), 2031–2041. http://dx.doi.org/10.1080/00036840210124559

Schneider, B. L., Wyse, A. E., & Keesler, V. (2007). *Is small really better? Testing some assumptions about high school size*. Brookings papers on Education Policy, pp. 15–47.

Silins, H., & Mulford, B. (2004). Schools as learning organisations—Effects on teacher leadership and student outcomes. *School Effectiveness and School Improvement, 15*(3–4), 443-466. http://dx.doi.org/10.1080/09243450512331383272

Stewart, E. A. (2003). School social bonds, school climate, and school misbehavior: A multilevel analysis. *Justice Quarterly, 20*(3), 575–604. http://dx.doi.org/10.1080/07418820300095621

Stewart, E. B. (2008). School structural characteristics, student effort, peer associations, and parental involvement the influence of school- and individual-level factors on academic achievement. *Education and Urban Society,40*(2), 179–204. doi:10.1177/0013124507304167

Stiefel, L., Berne, R., Iatarola, P., & Fruchter, N. (2000). High school size: Effects on budgets and performance in New York City. *Educational Evaluation and Policy Analysis,22*(1), 27–39.

Stiefel, L., Schwartz, A. L., & Ellen, I. G. (2006). Disentangling the racial test score gap: Probing the evidence in a large urban school district. *Journal of Policy Analysis and Management,26*(1), 7–30. doi:10.1002/pam.20225

Sun, L. T., Bradley, K. D., & Akers, K. (2012). A multilevel modelling approach to investigating factors impacting science achievement for secondary school students: PISA Hong Kong sample. *International Journal of Science Education,34*(14), 2107–2125. doi:10.1080/09500693.2012.708063

Tanner, K. C., & West, D. (2011). The effects of school size on academic outcomes. Retrieved on October 19, 2012, from http://sdpl.coe.uga.edu/research/SchoolSizeSDPL.pdf

Van der Vegt, A. L., den Blanken, M., & Hoogeveen, K. (2005). *Nationale scholierenmonitor: meting voorjaar 2005*. Utrecht: Sardes.

Vieno, A., Perkins, D. D., Smith, T. M., & Santinello, M. (2005). Democratic school climate and sense of community in school: A multilevel analysis. *American Journal of Community Psychology,36*(3–4), 327–341. doi:10.1007/s10464-005-8629-8

Watt, T. T. (2003). Are small schools and private schools better for adolescents' emotional adjustment? *Sociology of Education,76*(4), 344–367.

Wei, H. S., Williams, J. H., Chen, J. K., & Chang, H. Y. (2010). The effects of individual characteristics, teacher practice, and school organizational factors on students' bullying: A multilevel analysis of public middle schools. *Children and Youth Services Review,32*(1), 137–143. doi:10.1016/j.childyouth.2009.08.004

Weiss, C. C., Carolan, B. V., & Baker-Smith, E. C. (2010). Big school, small school: (Re)testing assumptions about high school size, school engagement and mathematics achievement. *Journal of Youth and Adolescence,39*(2), 163–176. doi:10.1007/s10964-009-9402-3

de Winter, M. (2003). *Niet te groot en niet te klein: effecten van schaalgrootte op het welbevinden van jongeren*. Utrecht: NIZW.

Wyse, A. E., Keesler, V., & Schneider, B. (2008). Assessing the effects of small school size on mathematics achievement: A propensity score-matching approach. *Teachers College Record,110*(9), 1879–1900.

Chapter 4
Quantitative Summary of Research Findings

Hans Luyten

This chapter presents a quantitative summary of research with regard to the effects of school size on student achievement and noncognitive outcomes (such as involvement, participation, social cohesion, safety, attendance, etc.). The noncognitive outcomes are widely considered as desirable in it, but are also often assumed to be conducive to high academic performance.

4.1 General Approach

The approach applied in this chapter yields an overall estimate of expected outcomes at a given school size. As such the approach can be considered a type of meta-analysis. However, common meta-analysis methods cannot be applied when dealing with research on the effects of school size. The main reason for this is that the relation between school size and outcomes is not always modeled as a clear-cut difference between small and large schools or as a straightforward linear relationship in studies that treat school size as a continuous variable.

Standard methods for conducting meta-analysis either assume a comparison between an experimental group and a comparison group or an effect measure that expresses a linear relationship. Outcomes from several studies are then standardized so that a weighted average effect can be computed (taking into account differences in sample sizes). The outcomes per study may be a standardized difference between groups (e.g., Cohen's d) or a statistic that describes the linear relation between an explanatory and a dependent variable (e.g., Fisher Z). Both kinds of measures can be converted to a common metric.

Many different forms beside a straightforward linear relationship (i.e., the smaller/larger the better) are hypothesized and reported in research on effects of school size, e.g., quadratic and log-linear. In a considerable number of studies, several different size categories are compared. The reason for this is that

H. Luyten et al., *School Size Effects Revisited*, SpringerBriefs in Education,
DOI: 10.1007/978-3-319-06814-5_4, © The Author(s) 2014

researchers want to take into account the possibility that it may be more appropriate to look for an optimal school size rather than to estimate a linear relationship between size and outcomes. Such a linear relationship would imply that the best results occur if schools would be either as large as a possible (e.g., one school for an entire district) or as small as possible (e.g., single class schools).

With regard to school size research, providing a quantitative summary of research findings is therefore quite complicated. Often more than just two school size categories are compared. In addition, the categories used vary between studies. In other cases the relation between school size and outcomes is modeled as a mathematical function (mostly linear, log-linear, or quadratic). The findings from these studies are not only difficult to compare to those that relate to comparisons between different school categories, but also the distinct mathematical functions cannot be converted to a common metric. When the effect of school size is modeled as a quadratic function, two distinct coefficients must be estimated (linear and quadratic), which precludes by definition converting the findings to a single metric.

As standard meta-analysis methods are not suitable when it comes to drawing up a quantitative summary of the research findings, another approach will be used. Based on the findings reported in the reviewed studies the "predicted" outcomes given a certain school size are calculated. To achieve comparability of the results only the scores on the outcome variables have been standardized to *z*-scores. There is no need to standardize the explanatory variable as well, because studies only have been included that use the same operationalization for school size (i.e., total number of students enrolled). Standardization of both the explanatory and the dependent variable is often applied in meta-analysis when the focus is on the relationship between two numerical variables. Often this is the only option available to render findings from different studies comparable, as the operationalizations of both dependent and independent variables tend to vary across studies. In such cases standardized regression coefficients may be the "raw material" processed in the meta-analysis. In the present case standardizing the independent variable is not required, but standardization of the outcomes is unavoidable, as the raw scores are incomparable across studies. Whatever the outcome variable relates to (student achievement, involvement, safety), the operationalization is bound to differ from one study to the next. The approach applied here reports for specific school sizes the average standardized outcomes over a number of studies. More details on this method are provided below as we illustrate more specifically how the "predicted" outcomes have been calculated for a couple of studies.

A potential risk of the approach relates to samples with strongly diverging ranges on the explanatory variable. Suppose that one is dealing with two samples. In the first sample, the school size ranges from very small (single class schools) to a total enrolment of 500 students and the average school size is 250. The second sample consists of schools with enrolments ranging from 500 to 1,000 students and the average school size is 750. If the effect of school size on achievement is

identical (e.g., achievement decreases one tenth of a standard deviation with a school size increase of 100 students enrolled), one would conclude that in both schools with 50 students and in schools with 550 students achievement is two tenths of a standard deviation above average. This interpretation might be correct, but it might also be mistaken. It is conceivable that the average achievement is much higher in the sample with smaller schools. In that case, the previous interpretation is clearly a mistake. It is therefore very important to be cautious in drawing conclusions from studies based on studies that vary strongly with regard to the ranges in school size. Note that similar risks apply to more commonly applied methods of meta-analysis.

4.2 Summarizing the Research Findings

Separate analyses are reported for student achievement and noncognitive outcomes. If an effect of school size for more than one measure of student achievement is reported (e.g., both language and mathematics), the average of these effects is reported in the summary. The same goes for noncognitive outcomes. In some studies, the effect of school size on a wide range of noncognitive outcomes (involvement, attendance, and safety) may be covered. Also in these cases the average effect is reported in the summary.

Findings will be reported separately for primary and secondary education. The main focus will be on the effect of school size on individual students. The key question addressed is to what extent student scores (cognitive or noncognitive) turn out to be relatively high or low given a certain school size. Student scores are standardized according to the well-known z-score transformation. First the mean is subtracted from each score and next the resulting difference is divided by the standard deviation. In another approach that is frequently applied, the result after subtraction is divided not by the standard deviation in student scores, but the standard deviation in school averages. The main argument for this approach is that school size, being a school level characteristic, can only have an impact on school means. Also when an analysis is based on data that are aggregated at the school level, it is hardly ever possible to estimate the effect of school size at the student level (unless information is available on the variation among student scores within schools). One highly important consequence of this approach is that it will inevitably yield larger estimates of school size effects. Only in the extreme situation where all variation in student scores is situated at the school level, (which would imply a complete absence of differences among students within schools) this approach yield the same estimate of a school size effect. However, as long as there is some variance between students within schools (which is always the case in real life), the school level variance (and therefore the standard deviation) is less than the total variance among students.

The argument outlined above may be illustrated with a simple example. Suppose that the standard deviation on a student achievement score equals 10

(therefore the variance is 100) and that the percentage of school level variance is 16.[1] This implies a variance equal to 16 and thus a standard deviation of 4 (square root of 16). Now let us assume that in large schools achievement scores are on average about 2 points lower than in small schools (for the moment, we will not deal with the question what counts as small and large school size). At the student level, this is a modest effect at best (one-fifth of a standard deviation), but if we compare the difference to the standard deviation among school means, the effect looks fairly impressive. In that case the difference between large and small schools equals half a standard deviation. Note that also such increases of the school size effect become even stronger as the percentage of school level variance decreases. In that case also the standard deviation among school means gets smaller, which will make the effect of school size appear to be larger. Especially for noncognitive outcomes differences between schools have often been reported to be quite modest.

In the authors' opinion, the most appropriate basis for expressing the school size effect is the total amount of variation (i.e., the standard deviation) among student scores. This puts the impact of school size in the right perspective. The impact is limited because it only affects school means. Most of the variation in student scores (both cognitive and cognitive) is situated within schools. This variation cannot be affected by changes in school size unless school size interacts with a student level variable (e.g., some studies have reported that the effect of school size is relatively for socioeconomically disadvantaged students). This natural limitation of the impact by school level characteristics should be clearly expressed in an assessment of the effects of school size. However, findings that are standardized by means of the school level standard deviation will be reported as well. Otherwise a substantial part of the available research would be discarded.

If one study covers two or more distinct samples (e.g., primary school students and secondary school students; or samples from different countries or regions) the outcomes per sample will be treated separately when the findings are reviewed. Thus, it is possible that a single study contributes more than one result when summarizing the findings.

Findings on the effect of school size are included in the summary if they meet the following two preconditions. First of all sufficient information needs to be provided for calculating the "predicted" outcomes at a given school size. Some author report only unstandardized regression coefficients without providing information on mean and standard deviations of the outcome variables. In such cases it is impossible to determine what the *standardized* outcome will be according to the regression model. In other cases only standardized regression coefficients are reported. In such cases one needs information on the mean and standard deviation of the explanatory variable (i.e., school size) in order to

[1] This is a realistic example. The total variance and percentage of variance at the school level are roughly the same for the standardized test taken in the final year of Dutch primary education (Cito eindtoets).

determine what the standardized outcomes will be for a given school size. The second precondition is that only findings are included if prior achievement has been controlled for. This is the case if the analysis is based on growth scores or if student achievement has been controlled for prior achievement. Note that controlling for cognitive aptitude (e.g., IQ measures) has only been counted as measures of prior achievement if students took the test at an earlier point in time. Some studies did control for cognitive aptitude that was measured during the same period that the outcome measures were collected. Findings from these studies have not been included in the quantitative summary.

4.3 School Size and Student Achievement in Primary Education

Out of the total number of studies on school size reviewed, five relate to its effect on individual student achievement in primary education and also meet the preconditions specified above. All five studies were conducted in the United States. Basic details about these studies are provided in Table 4.1.

In the studies by Archibald (2006), Holas and Huston (2012), and Maerten-Rivera et al. (2010) the relation between school size and student achievement is modeled as a linear function. Of these only Archibald (2006) reports a significant (and negative) effect of school size for both reading and mathematics. Maerten-Rivera et al. also report a negative effect, but a nonsignificant one. Holas and Huston (2012) only report that their analyses failed to reveal a significant relationship. For the quantitative summary of the research findings, it is assumed that they found a zero relationship. In the studies by Lee and Loeb (2000) and by Ready and Lee (2007) different categories of schools are compared. Lee and Loeb (2000) distinguish three categories (less than 400 students; 400–750 students, and over 750 students). Ready, and Lee (2007) distinguish five categories (less than 275 students; 275–400 students; 400–600 students; 600–800 students, and over 800 students). Lee and Loeb (2000) report significantly lower performance in the medium category (400–750 students) in comparison to the small category. Ready and Lee (2007) report significantly lower performance in the large schools category (>800 students) in comparison to the medium category (400–600) for reading in the first grade. For mathematics in the first grade they report a significantly higher performance in the small schools category (<275 students) in comparison to the medium category. No significant effects of school size were found in Kindergarten.

By taking a closer look at the findings reported by Archibald (2006) their implications become apparent in more detail. The reported standardized regression coefficients equal −0.03 and −0.07 for reading and mathematics, respectively. As the mean and standard deviation for school size are reported as well (see Table 4.1), any school size can be transformed into a z-score. The z-scores corresponding with school size ranging from 150 to 850 are displayed in Table 4.2. After that one only

Table 4.1 Studies on the effect of school size on individual achievement scores in primary education

	Archibald (2006)	Holas and Huston (2012)	Lee and Loeb (2000)	Maerten-Rivera et al. (2010)	Ready and Lee (2007)
Grade	3-6	5	6-8	5	Kindergarten, first grade
Location	US, Nevada	US, nation-wide	US, Chicago	US, Southeast	US, nation-wide
Outcomes	Reading, Mathematics	Reading, Mathematics	Mathematics	Science	Reading, mathematics
Sample size					
Schools	55	10	264	198	527
Students	7,000	804	22,599	23,854	7,740
School size					
Mean	547.8	540.0	500 (median)	798.1	500 (median)
Std. Dev.	137.4	260.0	not reported	330.9	not reported
Range	173–874	100–1,000 (approximation)	150–1,950	263–2,174	150–1,000 (approx.)

Table 4.2 Predicted z-scores for reading and mathematics per school size; based on findings report by Archibald (2006)

School size	z-score school size	Predicted z-scores		Average
		Reading	Mathematics	
150	−2.894	0.087	0.203	0.145
200	−2.531	0.076	0.177	0.127
250	−2.167	0.065	0.152	0.108
300	−1.803	0.054	0.126	0.090
350	−1.439	0.043	0.101	0.072
400	−1.075	0.032	0.075	0.054
450	−0.711	0.021	0.050	0.036
500	−0.348	0.010	0.024	0.017
550	0.016	0.000	−0.001	−0.001
600	−0.380	−0.011	−0.027	−0.019
650	0.744	−0.022	−0.052	−0.037
700	1.108	−0.033	−0.078	−0.055
750	1.472	−0.044	−0.103	−0.074
800	1.836	−0.055	−0.128	−0.092
850	2.199	−0.066	−0.154	−0.110

Table 4.3 Predicted z-scores for reading and mathematics per school size category; based on findings report by Lee and Loeb (2000)

School size categories	Standardized scores per school size category	
	School level	Student and school level combined
<400 students	0.054	0.026
400–750 students	−0.019	−0.009
>750 students	0.013	0.007

needs to multiply the *z*-scores with either −0.03 or −0.07 to arrive at the predicted *z*-scores for reading or mathematics. Table 4.2 shows the details that the Archibald findings imply in a primary school with 800 students and the reading scores on average are 0.055 of standard deviation below average. For mathematics this will be 0.128 of a standard deviation. The table also reports the average results across both subjects.

Lee and Loeb (2000) report differences in mathematics achievement between various school size categories after controlling for numerous confounding variables including prior achievement. The differences reported are standardized by dividing through the standard deviation among school averages. As both within school and between school variances are reported (Lee and Loeb 2000, p. 18), it is possible to rescale the reported differences relative to the total standard deviation in student achievement scores. Lee and Loeb (2000, p. 21) report that the math scores are on average 0.073 of a standard deviation higher in small schools vs. medium schools (less than 400 students versus 400–750 students). The advantage of small over large schools is more modest (0.041 and statistically not significant). Given the information provided in Lee and Loeb (2000) and assuming that the standardized average score must be equal to zero, it is possible to compute for each school size category the "predicted" average. Table 4.3 report two types of standardized scores. First the scores standardized relative to the standard deviation among school means and next the scores standardized relative to the total standard deviation in math scores (i.e., taking into account variation within and between schools). The table shows that the highest scores were found in the smallest schools. However, the differences are clearly more modest when they standardized relative to the standard deviation based on variation both within and between schools. The findings clearly suggest a curvilinear relationship between school size and achievement. Based on the standardized averages per category, a quadratic function has been estimated. This approach has also been applied to the findings reported by Ready and Lee (2007) and further on to findings from other studies that focus on differences between three or more school size categories.

Table 4.4 reports the main findings from all five studies on the school size effect in primary education based on student level findings. For each study, the predicted standardized achievement scores at student level are reported. All five studies report outcomes within the range from 200 to 850 students enrolled. For school sizes

Table 4.4 Predicted *student* achievement (standardized) per school size in primary education

Study	Archibald (2006)	Holas and Huston (2012)	Lee and Loeb (2000)	Maerten-Rivera et al. (2010)	Ready and Lee (2007)	Weighted average
Number of students	7,000	804	22,599	23,854	7,740	62,084
School size						
100		0.000				
150	0.145	0.000		0.105	0.028	
200	0.127	0.000	0.035	0.097	0.032	0.068
250	0.108	0.000	0.026	0.089	0.034	0.060
300	0.090	0.000	0.018	0.081	0.035	0.052
350	0.072	0.000	0.011	0.073	0.034	0.044
400	0.054	0.000	0.005	0.065	0.031	0.037
450	0.036	0.000	0.000	0.056	0.027	0.029
500	0.017	0.000	−0.004	0.048	0.021	0.022
550	−0.001	0.000	−0.007	0.040	0.014	0.014
600	−0.019	0.000	−0.009	0.032	0.005	0.008
650	−0.037	0.000	−0.010	0.024	−0.006	0.001
700	−0.055	0.000	−0.010	0.016	−0.018	−0.006
750	−0.074	0.000	−0.008	0.008	−0.032	−0.012
800	−0.092	0.000	−0.006	0.000	−0.048	−0.019
850	−0.110	0.000	−0.003	−0.008	−0.065	−0.025
900		0.000	0.002	−0.017	−0.084	
950		0.000	0.007	−0.025	−0.104	
1,000		0.000	0.013	−0.033	−0.126	
1,050				−0.041		
1,100				−0.049		
1150				−0.057		
1200				−0.065		
1250				−0.073		
1300				−0.081		
1350				−0.090		
1400				−0.098		
1450				−0.106		
1500				−0.114		
1550				−0.122		
1600				−0.130		
1650				−0.138		
1700				−0.146		
1750				−0.154		
1800				−0.163		
1850				−0.171		
1900				−0.179		

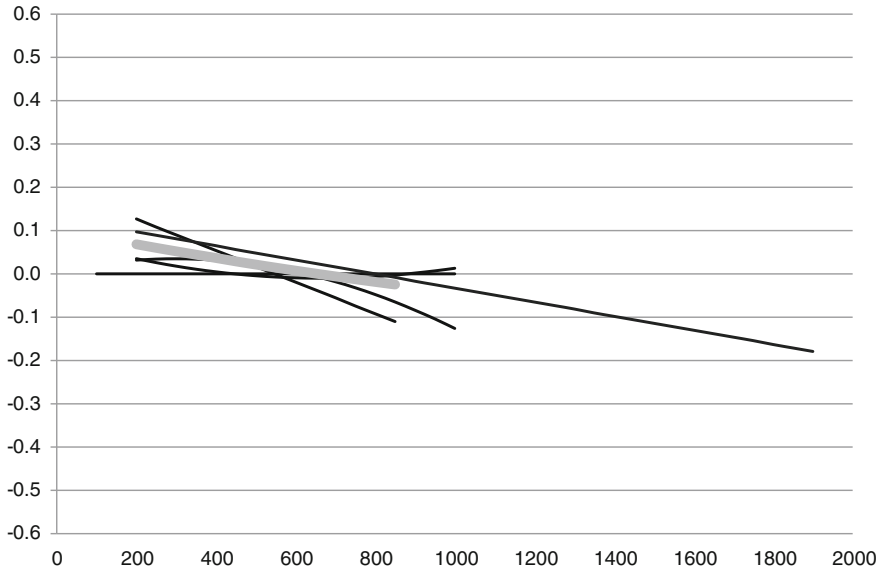

Fig. 4.1 Predicted STUDENT achievement per school size (primary education). The *thin black lines* represent the findings for a particular study; the *bold grey line* denotes the weighted average across studies

within this range a weighted average across all five studies has been calculated. Outcomes per study are weighted by the number of students.[2] Figure 4.1 provides a graphical display of the findings. On an average a slightly negative effect of school size on student achievement scores is detected. It should be noted, though, that the difference between student achievements in primary schools with 200 versus 850 students enrolled is still below one tenth of a standard deviation.

4.4 School Size and School Mean Achievement in Primary Education

It has already been mentioned that it also customary to standardize school size effects relative to the standard deviation among school means. This is the only

[2] In a meta-analysis based on effect sizes the results would be weighted by the inverse of the sampling variance. This weighting method cannot applied in the present case, as information on sampling variance was not reported for the predicted outcomes in any of the publications reviewed. Note that sampling variance is computed as the observed variance in the sample divided by the number of respondents. In the present case we can only take into account the number of respondents. De facto we assume that differences in variance between samples do not differ substantially.

Table 4.5 Studies on the effect of school size on school average achievement scores in primary education

	Archibald (2006)	Fernandez (2011)	Lee and Loeb (2000)	Maerten-Rivera et al. (2010)
Grade	3–6	3–10	6–8	5
Location	US, Nevada	US, Nevada	US, Chicago	US, Southeast
Outcomes	Reading, Mathematics	Reading, Mathematics	Mathematics	Science
Sample size				
Schools	55	252	264	198
School size				
Mean	547.8	1082,4	500 (median)	798.1
Std. Dev.	137.4	637.0	not reported	330.9
Range	173–874	205–3,311	150–1,950	263–2,174

option available when the analyses are based on aggregated school data. When multilevel analyses are conducted, it is possible to compute both types of standardized scores, provided that the necessary information on variance within and between schools on the outcome variable is reported. This is the case for three of the studies discussed in the previous section (Archibald 2006; Lee and Loeb 2000; Maerten-Rivera et al. 2010). See Table 4.5 for basic details on these studies. One additional study on school size and student achievement in primary education is included in Table 4.5 (Fernandez 2011). This study is based on aggregated school data. Like the other studies discussed so far, it relates to American schools (Nevada). The reported effect of school size on achievement is not significant and the standardized regression coefficient shows no noticeable deviation from zero. The study by Fernandez also includes high schools and middle schools, but the effects of school size are controlled for school type.

Appendix 4.1 presents the predicted standardized school means per school size for these studies. Figure 4.2 provides a graphical display of the findings. The figures in Appendix 4.1 also illustrate to what extent school size effects "increase" when the standardization is based on variation between school means. In the Archibald study the predicted standardized *student* scores range from 0.127 in schools with 200 students to −0.110 in schools with 800. The predicted standardized *school means* in the same study range from 0.296 to −0.257. Similar increases can be observed for the studies by Lee and Loeb (2000) and Maerten-Rivera et al. (2010). The impact of school size clearly appears to be more impressive if one compares the differences between large and small schools to the standard deviation of the school averages. Still, it is our opinion that the effects reported in Table 4.4 (i.e., impact on *student* scores) provide a more appropriate description of the impact of school size.

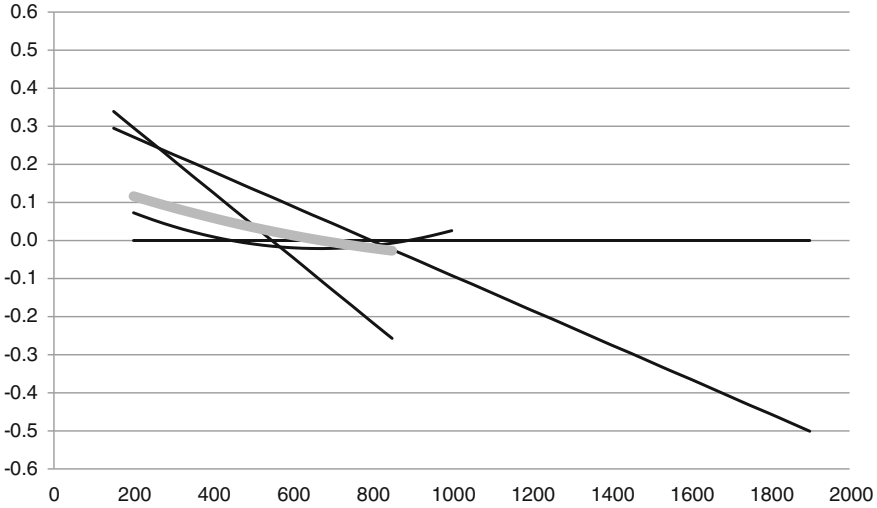

Fig. 4.2 Predicted MEAN SCHOOL achievement per school size (primary education). The *thin black lines* represent the findings for a particular study; the *bold grey line* denotes the weighted average across studies

4.5 School Size and Student Achievement in Secondary Education

Six studies have been found that related to the effect of school size on individual student achievement in secondary education and also meet our preconditions. Of these, five relate to secondary schools in the United States. The study by Ma and McIntyre (2005) deals with the situation in Canada (Alberta). Basic details are reported in Table 4.6. Except for the study by Ma and McIntyre (2005) the effect of school size is analyzed through comparison of different categories. However, there is little similarity in the categorizations applied. The number of categories range from 4 (Carolan 2012; Rumberger and Palardy 2005) to 8 (Lee and Smith 1997). See Table 4.6 for more details.

Most of the studies included in Table 4.6 report differences in student achievement between school categories. In those cases, a quadratic function has been estimated to describe the relation between school size and student achievement. This function is based on the standardized averages per category. There are two exceptions. The first one is the study by Luyten (1994), which only reports that no significant differences between categories were found. The other exception is the study by Ma and McIntyre (2005). Here a linear relation between school size and achievement is estimated, but the authors only report a significant interaction effect of taking math courses with school size on the mathematics post-test (the effect of taking math courses is weaker in larger schools; in other words: students that take math course get higher scores if they attend smaller schools). No main effect for

Table 4.6 Studies on the effect of school size on individual achievement scores in secondary education

	Carolan (2012)	Lee and Smith (1997)	Luyten (1994)	Ma and McIntyre (2005)	Rumberger and Palardy (2005)	Wyse et al. (2008)
Grade	12	12	8	12	12	12
Location	US, nationwide	US, nationwide	US, nationwide	Canada, Alberta	US, nationwide	US, nationwide
Outcomes	Math	Reading, math	Math	Math	Math, reading, science, history	Math
Sample size						
Schools	579	789	116	34	912	745
Students	9647	9812	4507	1518	14199	12853
School size						
Mean	1300 (median)	1050 (median)	not reported	612.8	900 (median)	1050 (median)
Std. Dev.	not reported	not reported	not reported	362.8	not reported	not reported
Range	400–1900 (approximation)	100–2400 (approx.)	150–1250 (approx.)	100–1300 (approx.)	300–2,100 (approx.)	200–2,200 (approx.)
School size categories	<600 600–999 1,000–1,599 > 1600	<300 301–600 601–900 901–1,200 1,201–1,500 1,501–1,800 1,801–2,100 >2,100	<240 240–360 360–500 500–1,000 >1,000	does not apply	<600 600–1,200 1200–1,800 >1,800	<400 400–799 800–1,199 1200–1,999 >2,000

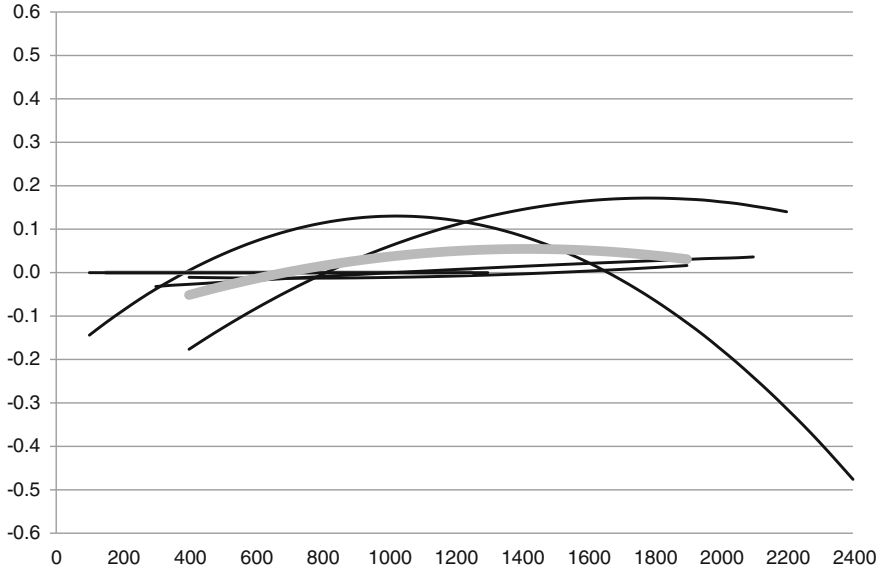

Fig. 4.3 Predicted STUDENT achievement per school size (secondary education). The *thin black lines* represent the findings for a particular study; the *bold grey line* denotes the weighted average across studies

school size on math achievement is reported. For this review it is assumed that the main school size effect is not statistically significant in this study. No further details are reported and for the summary of the research findings it is assumed that both the study Luyten (1994) and by Ma and McIntyre found a zero relationship.

Appendix 4.2 reports the predicted standardized achievement scores per school size in secondary education for individual student achievement. Weighted averages for school sizes within the range from 400 to 1,900 students enrolled are presented as well. Note that the studies by Luyten (1994) and Ma and McIntyre (2005) do not fully cover this range. The zero effects that are reported in these studies are assumed to extend beyond the exact ranges covered in these studies. In contrast to primary education, the findings suggest a curvilinear relation between school size and student achievement. The lowest scores are found in small secondary schools (−0.050). In schools with enrolments ranging from 1,200 to 1,600, the scores are at least one-tenth of a standard deviation higher. When schools get larger, the predicted scores decrease somewhat. Figure 4.3 provides a graphical display of the findings.

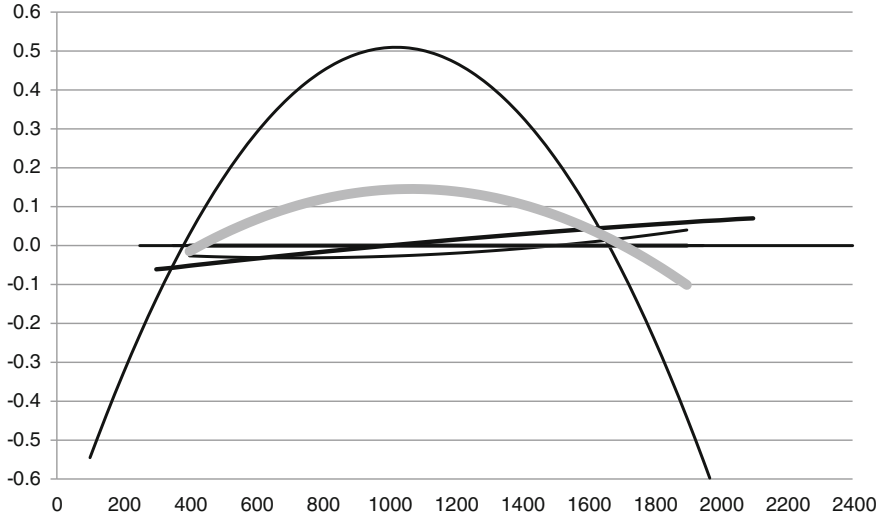

Fig. 4.4 Predicted MEAN SCHOOL achievement per school size (secondary education). The *thin black lines* represent the findings for a particular study; the *bold grey line* denotes the weighted average across studies

4.6 School Size and School Mean Achievement in Secondary Education

Appendix 4.3 reports the predicted standardized achievement scores per school size in secondary education for school mean achievement. For four out of the six studies included in Appendix 4.2, it was possible to calculate predicted standardized school means per school size. The study by Fernandez (2011), which makes use of aggregated school-level data (from the USA, Nevada) is included in Appendix 4.3. Again the findings reveal a curvilinear pattern, but now the lowest scores are found in the largest schools and the highest scores are found in schools with enrolments ranging from 900 to 1,250. This suggests a somewhat smaller optimum school size than suggested by the results based on individual achievement data. The findings from Appendix 4.3 are graphically displayed in Fig. 4.4.

4.7 School Size and Noncognitive Outcomes in Primary Education (Individual and School Means)

A wide range of outcome variables is subsumed under the label noncognitive outcomes. Still the number of studies on school size and noncognitive outcomes in primary education that report sufficient information to calculate the predicted outcomes per school size is quite limited, even though the requirements to be included in the quantitative summary are less stringent than for academic achievement. For

studies on noncognitive outcomes controlling for prior achievement was not considered necessary. Inclusion of socioeconomic background as a covariate in the analyses was deemed sufficient.

For the summary relating both to individual outcomes and school means five distinct studies are available. Of these, one relates exclusively to the effect of school size on individual outcomes (Holas and Huston 2012), two relate exclusively to school means (Durán-Narucki 2008; Lee and Loeb 2000) and two relate to both levels (Bonnet et al. 2009; Koth et al. 2008). See Table 4.7 for an overview of the studies on school size and noncognitive outcomes in primary education.

Four of the five studies listed in Table 4.7 report on American research. The other one relates to research in the Netherlands. In three studies, the effect of school size is modeled as a linear function (Durán-Narucki 2008; Holas and Huston 2012; Koth et al. 2008). In the other two studies, three categories are compared (Bonnet et al. 2009: <300, 301–500, >500; Lee and Loeb 2000: <400, 400–750, >750). When summarizing the findings, the results reported by Bonnet et al. (2009) have been rescored so that a high score denotes a positive situation (i.e., little peer victimization). These authors report significantly more victimization in the category of large schools (over 500 students). Lee and Loeb (2000) report significantly more positive teacher attitudes about responsibility for student learning in small schools (less than 400 students). Based on the standardized averages per category, a quadratic function has been estimated to denote the relation between school size and noncognitive outcomes in these two studies. Holas and Huston have analyzed the linear relation between school size and three noncognitive outcomes (student perceived self-competence, school involvement in grade 5, and in grade 6). Only the relation between size and involvement in grade 6 was found to be significant. The predicted scores presented in Appendix 4.4 denote the averages across these three outcomes. The study by Koth et al. (2008) focuses on achievement motivation and student-reported order and discipline. The relation between school size and order and discipline is not significant but they found a significantly negative relation between school size and achievement motivation. In Appendix 4.4, the averages across both outcomes are reported. Duran-Narucki focused on attendance and found significantly higher attendance in large schools (see Appendix 4.5). This is the only study on noncognitive outcomes in primary education that shows positive effects when schools are large.

The weighted average in Appendix 4.4 suggests a somewhat stronger effect of school size on noncognitive student outcomes in primary education as compared to achievement scores (see Table 4.4). The difference between primary schools with 200 versus 600 students is 0.13 standard deviation. With regard to student achievement scores, the difference between schools with 200 versus 600 students equals 0.076 standard deviation. Appendix 4.5 reports the predicted standardized school means per school size. The effect of school size looks stronger when standardized relative to standard deviation among school means. However, the standardization applied in Appendix 4.4 must be considered more appropriate. Graphic displays of the findings on the relation between school size and noncognitive outcomes in primary education are provided in Figs. 4.5 and 4.6.

Table 4.7 Studies on the effect of school size on individual and school mean noncognitive outcomes in primary education

	Bonnet et al. (2009)	Duran-Narucki (2008)	Holas and Huston (2012)	Koth et al. (2008)	Lee and Loeb (2000)
Grade	Kindergarten	All grades	6	5	6-8
Location	Netherlands	US, New York	US, nation-wide	US, Maryland	US, Chicago
Outcomes	Peer victimization (teacher reports)	Attendance	School attachment; Perceived self-competence (student reports)	Order and discipline; achievement motivation (student reports)	Teachers' attitudes about responsibility for student learning
Level	Student/school	School	Student	Student/school	School
Sample size					
Schools	23	95	10	37	264
Students	2003	not reported	804	2468	4495 teachers
School size					
Mean	400 (median)	711.8	540.0	488.4	500 (median)
Std. Dev.	not reported	328.3	260.0	146.7	not reported
Range	200–600 (approximation)	100-1,400 (approx.)	100–1,000 (approx.)	239–881	150-1,950

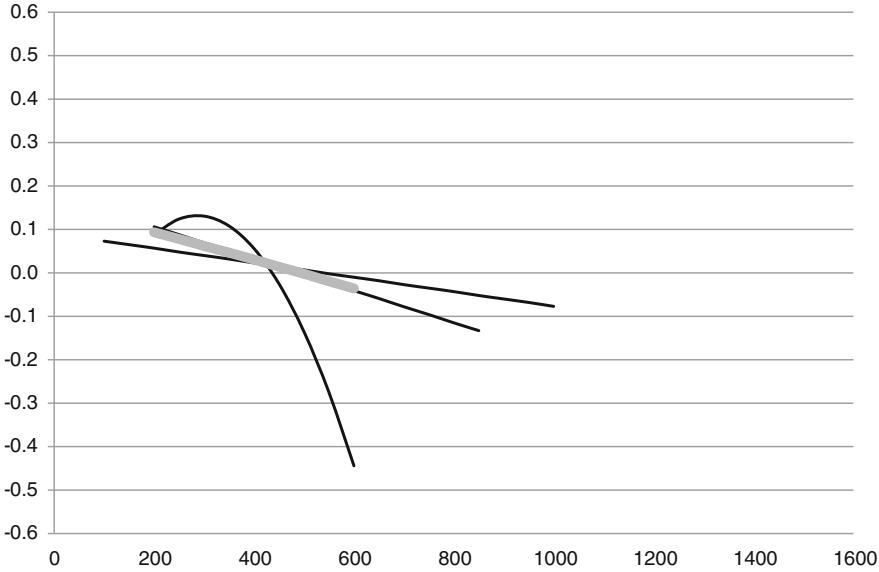

Fig. 4.5 Predicted non-cognitive STUDENT outcomes per school size (primary education). The *thin black lines* represent the findings for a particular study; the *bold grey line* denotes the weighted average across studies

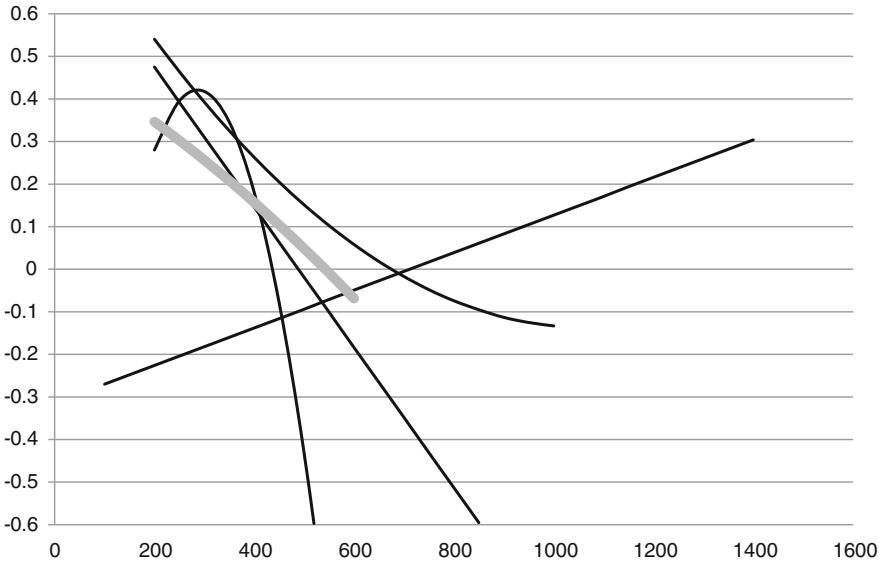

Fig. 4.6 Predicted non-cognitive MEAN SCHOOL outcomes per school size (primary education). The *thin black lines* represent the findings for a particular study; the *bold grey line* denotes the weighted average across studies

4.8 School Size and Noncognitive Outcomes in Secondary Education

A relatively large number of studies provide details on the predicted level of noncognitive outcomes per school size in secondary education. Table 4.8 provides basic information about these studies. The total number of *studies* is 19, but the study by Kirkpatrick Johnson et al. (2001) reports separate findings for middle schools (grades 7 and 8) and high schools (grades 7–12). As a result, the number of *samples* thus equals 20.

Twelve samples focus on the relation of school size with student outcomes and seventeen on the relation with school mean scores. Nine samples provide information on both student outcomes and school mean scores. Most research derives from the USA, but seven studies relate to other countries (two Israeli, two Dutch, the remaining three from Australia, Italy, and Taiwan). Many studies focus on the occurrence of incidents and other undesirable phenomena (such as harassment, disorder, theft, vandalism). All outcomes have been rescored in such a way that low scores denote a negative situation (e.g., high frequencies of vandalism and theft or low levels of safety or involvement). In most studies school size is modeled as a continuous variable. Only five studies make use of school size categories (Bowen et al. 2000; Chen 2008; Chen and Vazsonyi 2013; Dee et al. 2007; Rumberger and Palardy 2005). In the remaining 15 samples, the relation between school size and noncognitive outcomes is mostly modeled as a linear function, but in three cases (Gottfredson and DiPietro 2011; McNeely et al. 2002; Payne 2012) the researchers modelled it as a log-linear function (i.e., outcomes were regressed on the log of school size).

As shown in Table 4.8, many studies on noncognitive outcomes relate to multiple outcome measures. In these cases, the average effect of school size across the outcome measures involved has been computed. These are the outcomes reported in Appendices 6a–c and the corresponding figures.

4.9 School Size and Noncognitive Student Outcomes

Appendix 4.6a presents the findings for the American studies that focus on student outcomes. Appendix 4.6b reports the findings for the non-U.S. studies. The averages across studies (overall and broken down for American and non-U.S. samples) are reported in Appendix 4.6c. Graphic representations of the results are provided in the Figs. 4.7, 4.8 and 4.9.

For three out of the five American samples negative and significant effects on noncognitive outcomes are reported. The study by Gottfredson and DiPietro (2011) has come up with significantly positive effects. Kirkpatrick Johnson et al. (2001) report nonsignificant effects for their sample that focuses on students in middle schools. The strongest effect is reported in the study by Bowen et al.

Table 4.8 Studies on the effect of school size on noncognitive outcomes in secondary education

Study	Grades	Location	Outcomes	Level	Sample size		School size statistics		
					Schools	Students	Mean	Std. Dev.	Range
Attar-Schwartz (2009)	7–11	Israel	Sexual harassment	Student/school	327	16,604	557.6	332.3	42–1,643
Bowen et al. (2000)	6–12	US, nationwide	School satisfaction, teacher support, school safety	Student	39	945	689.0	–	70–1,393
Chen (2008)	7–12	US, nationwide	Criminal incidents	School	712	–	median >999	–	200–1,250 (approx.)
Chen and Vazsonyi (2013)	mean age: 16.6	US, nationwide	Problem behaviors (dishonesty to parents, selling drugs, using alcohol)	Student/school	85	9163	median: ±1000	–	200–1,200 (approx.)
(Chen 2008)	8	US, New York City	School disorder (crime, minor crime, noncrime)	School	212	–	959.7	493.4	164–2,262
Dee et al. (2007)	10	US, nationwide	Parental involvement	Student	±390	8,197	median: ±1700	–	200–2,700 (approx.)
Gottfredson and DiPietro (2011)	mean age: 14.1	US, nationwide	Safety: personal and property victimization	Student/school	253	13,597	792.0	478.6	97–2,912
Haller (1992)	10 and 12	US, nationwide	Student indiscipline (truancy, vandalism/theft), principal/student perceptions, student self-reports	School	558	–	962.7	1218.7	100–2,800 (approx.)
Khoury-Kassabri et al. (2004)	7–11	Israel	Student victimization (serious physical, moderate physical, verbal-social, threats)	Student/school	162	10400	505.3	297.5	100–1,100 (approx.)

(continued)

Table 4.8 (continued)

Study	Grades	Location	Outcomes	Level	Sample size		School size statistics		
					Schools	Students	Mean	Std. Dev.	Range
Kirkpatrick Johnson et al. (2001)	7–8 (middle) 7–12 (high)	US, nationwide	Student attachment to school (feelings); academic engagement (behavior)	Student/ school	45 (mid) 64 (hi)	2,482 (mid) 8,104 (hi)	477 (mid) 1,147 (hi)	234 (mid) 716 (hi)	150–800; 100–2,600 (approx.)
Klein & Cornell (2010)	9	US, Virginia	Victimization: bullying, threats and physical attack (student self-reports, school records)	School	290	7431	1210	690	100–2600 (approx.)
McNeely et al. (2002)	7–12	US, nationwide	Student connectedness to school	School	127	75515	642	765	42–5,422
Mooij, Smeets and De Wit (2011)	7–12	Netherlands	Safety: ignoring, excluding, threatening, intimidating, blackmailing, spreading rumours	Student/ school	104	26162	926	514	21–2,336
Payne (2012)	mean age: 14.1	US, nationwide	Communal school organization (supportive and collaborative faculty/ staff relations; common goals and norms)	School	253	–	792.0	478.6	100–1,700 (approx.)
Rumberger and Palardy (2005)	12	US, nationwide	Attrition (dropout and transfer rates to another school)	School	912	14,199	median: ±900	–	300–2,100 (approx.)

(continued)

Table 4.8 (continued)

Study	Grades	Location	Outcomes	Level	Sample size		School size statistics		
					Schools	Students	Mean	Std. Dev.	Range
Silins and Mulford (2004)	10	Australia	Student participation, student engagement, organisational learning, teacher leadership	School	96	3,500	631.9	283.2	100–1,200 (approx.)
Van der Vegte tal. (2005)	7–12	Netherlands	Safety, well-being	Student	51	5206	785.1	527.2	100–2,200 (approx.)
Vieno et al. (2005)	6, 8 & 10	Italy	Sense of community (belonging) at school	Student/ school	134	4733	480.1	303.6	52–1,509
Wei et al. (2010)	7–9	Taiwan	Bullying behavior: physical, verbal (self-reported)	Student/ school	12	1172	1567.6	988.7	300–2,800 (approx.)

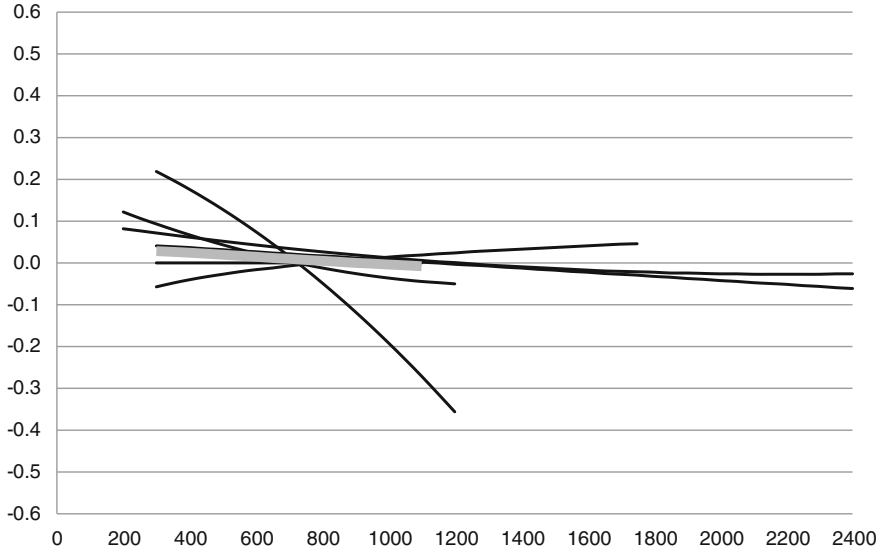

Fig. 4.7 Predicted non-cognitive STUDENT outcomes per school size (secondary education; American Studies). The *thin black lines* represent the findings for a particular study; the *bold grey line* denotes the weighted average across studies

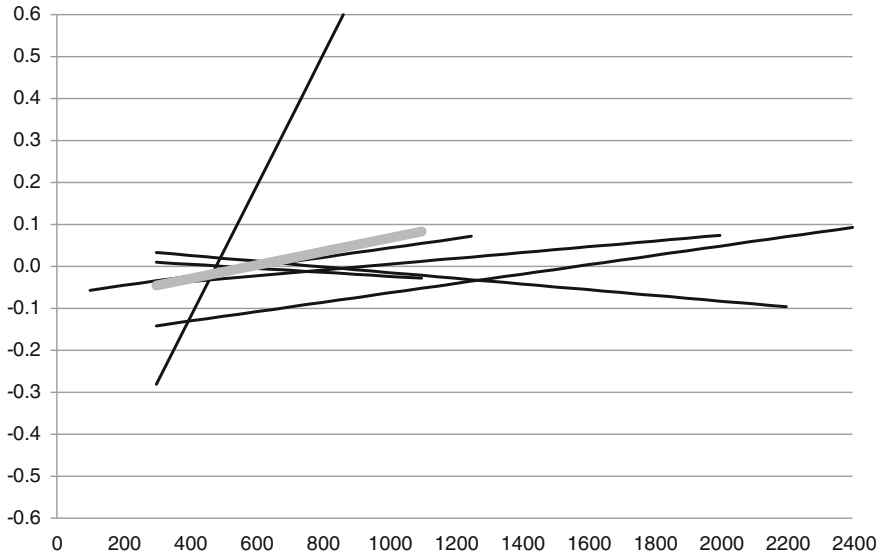

Fig. 4.8 Predicted non-cognitive STUDENT outcomes per school size (secondary education; non-U.S. studies). The *thin black lines* represent the findings for a particular study; the *bold grey line* denotes the weighted average across studies

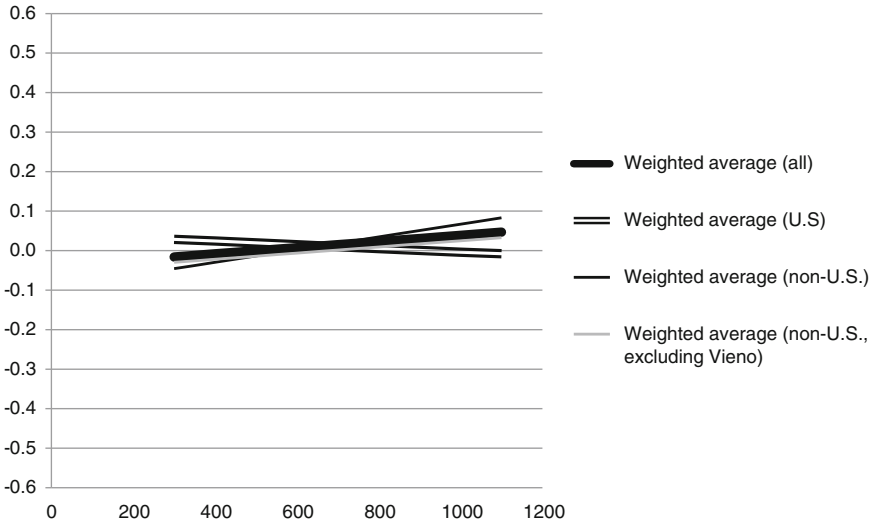

Fig. 4.9 Average outcomes non-cognitive STUDENT scores per school size (secondary education)

(2000), which reports a difference of about half a standard deviation between the smallest and the largest schools. School size ranges in this study from less than 100 students to nearly 1,400. The outcome measures relate to school satisfaction, safety, and teacher support.

Whereas the American findings mostly show negative effects of large school size on noncognitive student outcomes in secondary education, research conducted outside the U.S. fails to confirm this picture. Appendix 4.6b presents the results from six studies conducted outside the U.S. Of these, three show a negative effect of large school size, but the other three show a positive effect. Two of the negative effects are statistically significant (Attar-Schwartz 2009; Van der Vegt et al. 2005). Only one of the reported positive effects is significant (Mooij et al. 2011). All of these three studies relate to various aspects of school safety. Two of these studies were conducted in the Netherlands. Both reports show significant effects, but in different directions. The finding reported by Vieno et al. (2005) for Italy deserves special mention. The effect in this study appears to be particularly strong, without reaching statistical significance. Perhaps the strong effect is due to over-fitting, as the number of explanatory variables at the school level is quite large relative to the number of schools.

The general picture on the relation between school size and noncognitive outcomes at the student level across all twelve samples is provided in Appendix 4.6c and Fig. 4.9. The overall effects of school size on noncognitive student outcomes appear to be quite modest, but findings from the U.S. versus outside the U.S. contradict each other. The average effect in American studies is slightly negative, whereas studies form other countries (Israel, Italy, the Netherlands, and

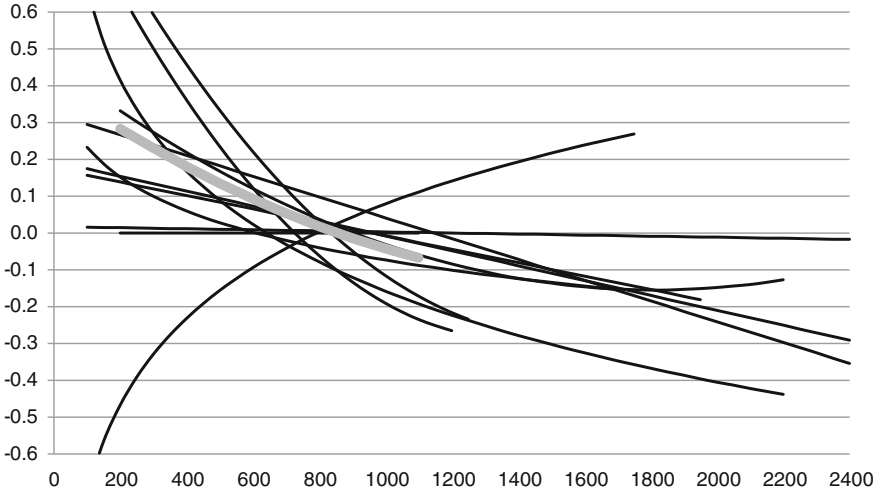

Fig. 4.10 Predicted non-cognitive SCHOOL MEAN outcomes per school size (secondary education; American Studies). The *thin black lines* represent the findings for a particular study; the *bold grey line* denotes the weighted average across studies

Taiwan) show on average a positive effect of school size. Even when the findings from the study by Vieno et al. (2005) are excluded from the summary, the effect of school size remains positive. However, the effect becomes considerably smaller in that case. School size effects on noncognitive student outcomes must be described as small. The difference between predicted scores in schools with 300 versus 1,100 students is about 0.06 of standard deviation (positive or negative). The findings that relate to the U.S. suggest a negative effect of large school size, but this average effect is even smaller than the positive effects found in other countries.

4.10 School Size and Noncognitive School Mean Scores

The findings that relate to the relation between school size and standardized school mean scores largely replicate the findings on student outcomes. The main difference is that the effect on school mean scores *appears* to be stronger. This is basically a statistical artifact as the variation in school means is bound to be smaller than the variation between student scores. Again we see negative, but relatively small effects of large school size in the USA, while a reverse picture emerges from non-U.S. research. More details are provided in Appendices 7a–c and Figs. 4.10, 4.11, and 4.12 provide graphic illustrations of the trends described.

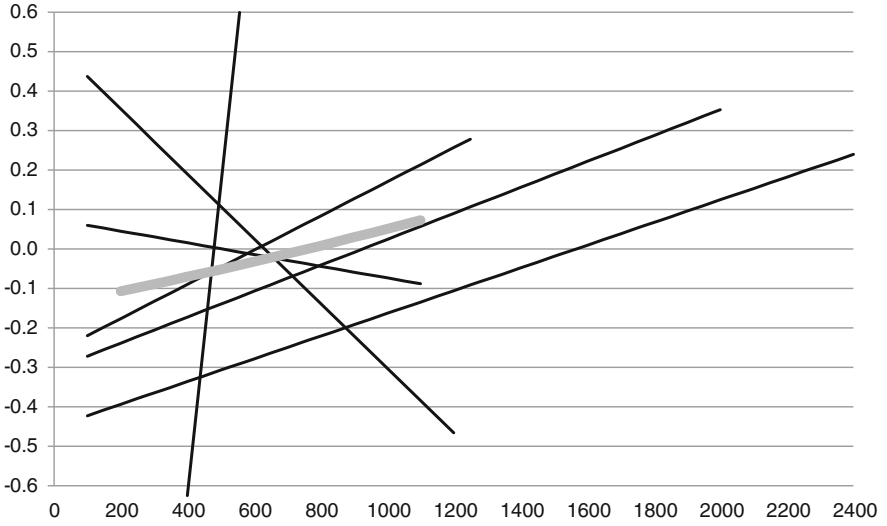

Fig. 4.11 Predicted non-cognitive SCHOOL MEAN outcomes per school size (secondary education; non-U.S. studies). The *thin black lines* represent the findings for a particular study; the *bold grey line* denotes the weighted average across studies

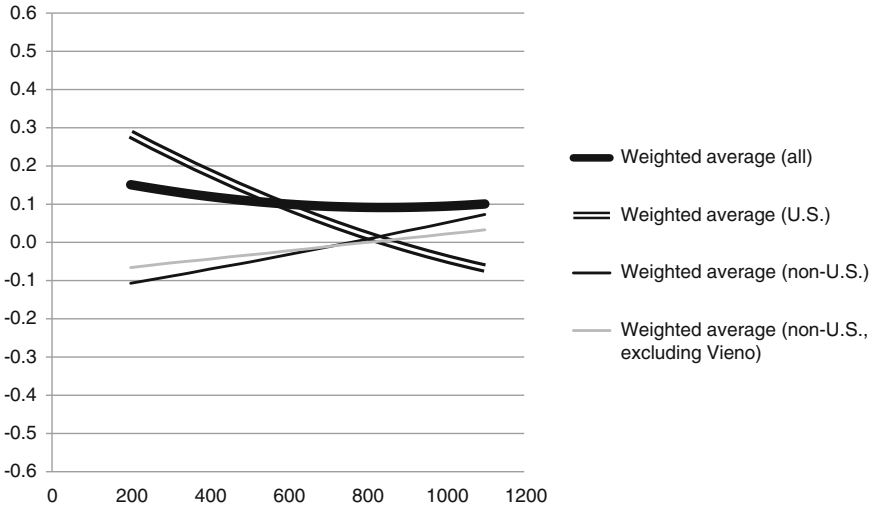

Fig. 4.12 Average outcomes non-cognitive STUDENT scores per school size (secondary education)

4.11 Conclusion

The research synthesis presented in this chapter was aimed at a precise specification of the relationship between school size and outcomes (both cognitive and noncognitive) in primary and secondary education. The predicted level of standardized outcomes given a certain school size was calculated for dozens of samples, based on the information provided in reports on the effects of school size. The discussion of the findings will focus on results related to outcomes that are standardized through division by the standard deviation in student scores. The alternative (division by the standard deviation in school means) is considered as less appropriate. It is bound to produce results that appear to reveal stronger effects of school size, which is confirmed in the present report. However, this approach tends to obscure that school size is unlikely to affect variation in student outcomes within schools, whereas the bulk of the variation in student scores (cognitive and noncognitive) is situated within schools.

On an average the review shows a slightly negative relation in primary education between school sizes both for cognitive and noncognitive outcomes. It should be noted that this finding is almost exclusively based on American research. The difference in predicted scores between very small and large schools is less than one tenth of a standard deviation for cognitive outcomes and somewhat larger (0.13 standard deviation) for noncognitive outcomes. Taken into account that the difference between the smallest and the largest schools amount at least to two standard deviations, it is clear that the effect of school size in terms of a standardized effect size (e.g., Cohen's d) must be very modest. For noncognitive outcomes, it may still exceed the (very modest) value of 0.05, but for cognitive outcomes the effect is even weaker.

For cognitive outcomes secondary education, a curvilinear pattern emerged from the studies reviewed. The highest scores appear to occur in schools with over 1,200 students but less than 1,600 students. In larger schools, lower scores are found, but the lowest scores are predicted for schools with less than 700 students. The difference between the lowest scoring schools (400 students) and the highest scoring (1,350–1,500 students) is just over one-tenth of a standard deviation. Because the relation between school size and outcomes does not always fit into a linear pattern, it is difficult to express it in more current metrics like Cohen's d, or a correlation coefficient. The difference between the highest scoring schools (i.e., medium to large) and small schools is probably less than one tenth of a standard deviation, which would commonly be considered a small effect (i.e., Cohen's $d < 0.20$). This assessment is based on the supposition that the difference in size between very small and medium to large schools (approximately 1,000 students) accounts for atleast one standard deviation.[3] The findings on cognitive outcomes are exclusively based on research conducted in the U.S.

[3] If the standard deviation in school size is 5,00 instead of 1,000, a difference of 0.10 would imply an effect size of 0.05.

With regard to research on the relation between school size and noncognitive outcomes in secondary education a large part of the results relate to studies from other countries as well. Interestingly, clearly opposite trends are apparent in American studies versus studies from other countries. Across all studies the trend is slightly in favor of large schools. The difference between small secondary schools (300 students) and large ones (1,100 students) amounts to 0.06 standard deviation, but for American studies the trend is reversed. Small schools show more favorable scores, although the difference between small and large American schools turns out to be very modest (0.04 standard deviation). The effect of school size in non-U.S. studies is somewhat stronger and reversed (showing more positive scores in large schools).

Appendix 1: Predicted *School Mean* Achievement (Standardized) Per School Size in Primary Education

	Archibald (2006)	Fernandez (2011)	Lee and Loeb (2000)	(Maerten-Rivera et al. 2010)	Weighted average
Number of schools	55	252	264	198	769
School size					
100					
150	−0.339			0.295	
200	0.296	0.000	0.073	0.272	0.116
250	0.254	0.000	0.054	0.249	0.101
300	0.211	0.000	0.037	0.226	0.086
350	0.168	0.000	0.022	0.204	0.072
400	0.126	0.000	0.010	0.181	0.059
450	0.083	0.000	−0.001	0.158	0.046
500	0.041	0.000	−0.009	0.135	0.035
550	−0.002	0.000	−0.015	0.113	0.024
600	−0.044	0.000	−0.019	0.090	0.014
650	−0.087	0.000	−0.021	0.067	0.004
700	−0.130	0.000	−0.021	0.045	−0.005
750	−0.172	0.000	−0.018	0.022	−0.013
800	−0.215	0.000	−0.014	−0.001	−0.020
850	−0.257	0.000	−0.007	−0.024	−0.027
900		0.000	0.002	−0.046	
950		0.000	0.013	−0.069	
1,000		0.000	0.026	−0.092	
1,050		0.000		−0.114	
1,100		0.000		−0.137	
1,150		0.000		−0.160	

(continued)

(continued)

	Archibald (2006)	Fernandez (2011)	Lee and Loeb (2000)	(Maerten-Rivera et al. 2010)	Weighted average
1,200		0.000		−0.183	
1,250		0.000		−0.205	
1,300		0.000		−0.228	
1,350		0.000		−0.251	
1,400		0.000		−0.274	
1,450		0.000		−0.296	
1,500		0.000		− 0.319	
1,550		0.000		− 0.342	
1,600		0.000		− 0.364	
1,650		0.000		− 0.387	
1,700		0.000		− 0.41	
1,750		0.000		− 0.433	
1,800		0.000		− 0.455	
1,850		0.000		− 0.478	
1,900		0.000		− 0.501	

Appendix 2: Predicted *Student* Achievement (Standardized) Per School Size in Secondary Education

	Carolan (2012)	Lee and Smith (1997)	Luyten (1994)	Ma and McIntyre (2005)	Rumberger and Palardy (2005)	Wyse et al. (2008)	Weighted average
N (students)	9,647	9,812	4,507	1,518	14,199	12,853	54,134
School size							
100		−0.144		0.000			
150		−0.115	0.000	0.000			
200		−0.088	0.000	0.000		−0.284	
250		−0.062	0.000	0.000		−0.256	
300		−0.038	0.000	0.000	−0.032	−0.228	
350		−0.016	0.000	0.000	−0.029	−0.202	
400	−0.011	0.005	0.000	0.000	−0.027	−0.176	−0.051
450	−0.011	0.024	0.000	0.000	−0.025	−0.152	−0.041
500	−0.012	0.042	0.000	0.000	−0.022	−0.128	−0.032
550	−0.012	0.058	0.000	0.000	−0.020	−0.105	−0.022
600	−0.013	0.072	0.000	0.000	−0.018	−0.083	−0.014
650	−0.013	0.085	0.000	0.000	−0.015	−0.062	−0.006
700	−0.013	0.096	0.000	0.000	−0.013	−0.042	0.002
750	−0.013	0.106	0.000	0.000	−0.011	−0.023	0.009
800	−0.013	0.114	0.000	0.000	−0.009	−0.004	0.015
850	−0.012	0.120	0.000	0.000	−0.007	0.013	0.021
900	−0.012	0.125	0.000	0.000	−0.005	0.029	0.027

(continued)

(continued)

	Carolan (2012)	Lee and Smith (1997)	Luyten (1994)	Ma and McIntyre (2005)	Rumberger and Palardy (2005)	Wyse et al. (2008)	Weighted average
950	−0.012	0.128	0.000	0.000	−0.003	0.045	0.032
1,000	−0.011	0.130	0.000	0.000	−0.001	0.060	0.037
1,050	−0.011	0.130	0.000	0.000	0.001	0.074	0.041
1,100	−0.010	0.128	0.000	0.000	0.003	0.086	0.044
1,150	−0.009	0.125	0.000	0.000	0.005	0.098	0.047
1,200	−0.008	0.120	0.000	0.000	0.007	0.109	0.050
1,250	−0.007	0.114	0.000	0.000	0.009	0.120	0.052
1,300	−0.006	0.106		0.000	0.011	0.129	0.053
1,350	−0.004	0.096			0.013	0.137	0.054
1,400	−0.003	0.085			0.014	0.145	0.055
1,450	−0.002	0.072			0.016	0.151	0.054
1,500	0.000	0.058			0.018	0.157	0.054
1,550	0.002	0.042			0.019	0.162	0.053
1,600	0.003	0.024			0.021	0.165	0.051
1,650	0.005	0.005			0.023	0.168	0.049
1,700	0.007	−0.016			0.024	0.170	0.046
1,750	0.009	−0.038			0.026	0.171	0.043
1,800	0.012	−0.063			0.027	0.171	0.040
1,850	0.014	−0.088			0.029	0.171	0.036
1,900	0.016	−0.115			0.030	0.169	0.031
1,950		−0.144			0.032	0.167	
2,000		−0.175			0.033	0.163	
2,050		−0.207			0.034	0.159	
2,100		−0.241			0.036	0.153	
2,150		−0.276				0.147	
2,200		− 0.313				0.140	
2,250		− 0.351					
2,300		− 0.391					
2,350		− 0.433					
2,400		− 0.476					

Appendix 3: Predicted *School Mean* Achievement (Standardized) Per School Size in Secondary Education

	Carolan (2012)	Fernandez (2011)	Lee and Smith (1997)	Luyten (1994)	Ma and McIntyre (2005)	Rumberger and Palardy (2005)	Weighted average
N (schools)	579	252	789	116	34	912	2,648
School size							
100			− 0.545		0.000		
150			− 0.434	0.000	0.000		
200			− 0.329	0.000	0.000		
250		0.000	−0.230	0.000	0.000		

(continued)

(continued)

	Carolan (2012)	Fernandez (2011)	Lee and Smith (1997)	Luyten (1994)	Ma and McIntyre (2005)	Rumberger and Palardy (2005)	Weighted average
300		0.000	−0.137	0.000	0.000	−0.061	
350		0.000	−0.051	0.000	0.000	−0.057	
400	−0.027	0.000	0.029	0.000	0.000	−0.052	−0.015
450	−0.028	0.000	0.103	0.000	0.000	−0.047	0.008
500	−0.029	0.000	0.171	0.000	0.000	−0.042	0.030
550	−0.030	0.000	0.233	0.000	0.000	−0.038	0.049
600	−0.031	0.000	0.288	0.000	0.000	−0.033	0.067
650	−0.031	0.000	−0.337	0.000	0.000	−0.029	0.083
700	−0.032	0.000	−0.380	0.000	0.000	−0.025	0.096
750	−0.032	0.000	−0.417	0.000	0.000	−0.020	0.109
800	−0.031	0.000	−0.448	0.000	0.000	−0.016	0.120
850	−0.031	0.000	−0.473	0.000	0.000	−0.012	0.128
900	−0.030	0.000	−0.491	0.000	0.000	−0.008	0.135
950	−0.029	0.000	−0.503	0.000	0.000	−0.004	0.140
1,000	−0.028	0.000	−0.509	0.000	0.000	0.000	0.144
1,050	−0.026	0.000	−0.509	0.000	0.000	0.004	0.145
1,100	−0.024	0.000	−0.502	0.000	0.000	0.008	0.145
1,150	−0.022	0.000	−0.490	0.000	0.000	0.012	0.143
1,200	−0.020	0.000	−0.471	0.000	0.000	0.015	0.139
1,250	−0.017	0.000	−0.446	0.000	0.000	0.019	0.134
1,300	−0.014	0.000	−0.415		0.000	0.022	0.127
1,350	−0.011	0.000	−0.377			0.026	0.117
1,400	−0.008	0.000	−0.334			0.029	0.106
1,450	−0.004	0.000	0.284			0.033	0.094
1,500	0.000	0.000	0.228			0.036	0.079
1,550	0.004	0.000	0.166			0.039	0.063
1,600	0.008	0.000	0.098			0.042	0.045
1,650	0.013	0.000	0.023			0.045	0.025
1,700	0.018	0.000	−0.057			0.048	0.003
1,750	0.023	0.000	−0.144			0.051	−0.020
1,800	0.029	0.000	−0.237			0.054	−0.045
1,850	0.034	0.000	−0.336			0.057	−0.072
1,900	0.040	0.000	−0.442			0.060	−0.101
1,950		0.000	−0.553			0.063	
2,000		0.000	−0.671			0.065	
2,050		0.000	−0.795			0.068	
2,100		0.000	−0.925			0.070	
2,150		0.000	−1.062				
2,200		0.000	−1.204				
2,250		0.000	−1.353				
2,300		0.000	−1.508				
2,350		0.000					
2,400		0.000					

Appendix 4: Predicted Noncognitive *Student* Outcomes (Standardized) Per School Size in Primary Education

	Bonnet et al. (2000)	Holas and Huston (2012)	Koth et al. (2008)	Weighted average
N (students)	2,003	855	2,468	5,326
School size				
100		0.073		
150		0.065		
200	0.088	0.057	0.106	0.093
250	0.124	0.048	0.088	0.078
300	0.131	0.040	0.069	0.062
350	0.108	0.032	0.051	0.047
400	0.056	0.023	0.033	0.031
450	−0.025	0.015	0.014	0.014
500	−0.136	0.007	−0.004	−0.002
550	−0.275	−0.002	−0.023	−0.019
600	−0.444	−0.010	−0.041	−0.036
650		−0.018	−0.059	
700		−0.027	−0.078	
750		−0.035	−0.096	
800		−0.043	−0.115	
850		−0.052	−0.133	
900		−0.060		
950		−0.068		
1,000		−0.077		
1,050				
1,100				
1,150				
1,200				
1,250				

Appendix 5: Predicted Noncognitive *School Mean* Outcomes (Standardized) Per School Size in Primary Education

	Bonnet et al. (2009)	Durán-Narucki (2008)	Koth et al. (2008)	Lee and Loeb (2000)	Weighted average
N (schools)	23	95	37	264	419
School size					
100		−0.270			
150		−0.248			
200	0.280	−0.226	−0.475	−0.540	−0.346
250	−0.396	−0.204	−0.392	−0.464	−0.302
300	−0.418	−0.182	−0.310	−0.392	0.256
350	−0.346	−0.160	0.228	−0.325	0.208
400	0.180	−0.138	0.146	0.263	0.157
450	−0.079	−0.116	0.063	0.205	0.104
500	−0.433	−0.094	−0.019	0.151	0.049
550	−0.879	−0.071	−0.101	0.102	−0.009
600	−1.420	−0.049	−0.184	0.058	−0.069
650		−0.027	−0.266	0.018	
700		−0.005	−0.348	−0.017	
750		0.017	−0.431	−0.048	
800		0.039	−0.513	−0.074	
850		0.061	−0.595	−0.095	
900		0.083		−0.113	
950		0.105		−0.125	
1,000		0.127		−0.133	
1,050		0.149			
1,100		0.171			
1,150		0.194			
1,200		0.216			
1,250		0.238			
1,300		0.260			
1,350		0.282			
1,400		−0.304			

Appendix 6a: Predicted Noncognitive *Student* Outcomes Per School Size in Secondary Education; American Studies

	Bowen et al. (2000)	Chen and Vazsonyi (2013)	Dee et al. (2007)	Gottfredson and DiPietro (2011)	Kirkpatrick Johnson et al. (2001); middle schools	Kirkpatrick Johnson et al. (2001); high schools
N (students)	945	9,163	8,197	13,597	2,482	8,104
School size						
100				−0.121		0.051
150				−0.097	0.000	0.049
200		0.122	0.082	−0.081	0.000	0.046
250		0.107	0.077	−0.067	0.000	0.044
300	0.219	0.093	0.072	−0.057	0.000	0.041
350	0.198	0.080	0.067	−0.048	0.000	0.039
400	0.176	0.067	0.062	−0.04	0.000	0.037
450	0.153	0.055	0.057	−0.033	0.000	0.034
500	0.128	0.044	0.052	−0.027	0.000	0.032
550	0.102	0.033	0.048	−0.021	0.000	0.029
600	0.075	0.022	0.043	−0.016	0.000	0.027
650	0.046	0.013	0.039	−0.012	0.000	0.024
700	0.016	0.004	0.035	−0.007	0.000	0.022
750	−0.015	−0.004	0.031	−0.003	0.000	0.019
800	−0.048	−0.012	0.027	0.001	0.000	0.017
850	−0.081	−0.019	0.023	0.004		0.015
900	−0.117	−0.025	0.019	0.007		0.012
950	−0.153	−0.031	0.016	0.011		0.010
1,000	−0.191	−0.036	0.013	0.014		0.007
1,050	−0.230	−0.041	0.009	0.017		0.005
1,100	−0.271	−0.044	0.006	0.019		0.002
1,150	−0.313	−0.047	0.003	0.022		0.000
1,200	−0.356	−0.050	0.001	0.024		−0.003
1,250			−0.002	0.027		−0.005
1,300			−0.005	0.029		−0.007
1,350			−0.007	0.031		−0.010
1,400			−0.009	0.033		−0.012
1,450			−0.011	0.035		−0.015
1,500			−0.013	0.037		−0.017
1,550			−0.015	0.039		−0.020
1,600			−0.017	0.041		−0.022
1,650			−0.019	0.043		−0.025
1,700			−0.020	0.045		−0.027
1,750			−0.021	0.046		−0.029
1,800			−0.022			−0.032

(continued)

(continued)

	Bowen et al. (2000)	Chen and Vazsonyi (2013)	Dee et al. (2007)	Gottfredson and DiPietro (2011)	Kirkpatrick Johnson et al. (2001); middle schools	Kirkpatrick Johnson et al. (2001); high schools
1,850			−0.024			−0.034
1,900			−0.024			−0.037
1,950			−0.025			−0.039
2,000			−0.026			−0.042
2,050			−0.026			−0.044
2,100			−0.027			−0.047
2,150			−0.027			−0.049
2,200			−0.027			−0.051
2,250			−0.027			−0.054
2,300			−0.027			−0.056
2,350			−0.026			−0.059
2,400			−0.026			−0.061

Appendix 6b: Predicted Noncognitive *Student* Outcomes Per School Size in Secondary Education; Studies Outside the U.S

	Attar-Schwarz (2009)	Khoury-Kassabri et al. (2004)	Mooij et al. (2011)	Van der Vegt et al. (2005)	Vieno et al. (2005)	Wei et al. (2010)
N (students)	16,604	1 − 0.400	26,162	5,206	4,733	1,172
School size						
100	−0.057	0.019	−0.057	0.047	−0.593	−0.164
150	−0.051	0.017	−0.054	0.043	−0.515	−0.158
200	−0.045	0.015	−0.050	0.040	−0.437	−0.153
250	−0.040	0.012	−0.047	0.036	−0.359	−0.147
300	−0.034	0.010	−0.043	0.033	−0.281	−0.142
350	−0.029	0.007	−0.040	0.030	−0.203	−0.136
400	−0.023	0.005	−0.036	0.026	−0.125	−0.130
450	−0.018	0.003	−0.033	0.023	−0.047	−0.125
500	−0.012	0.000	−0.029	0.019	0.031	−0.119
550	−0.006	−0.002	−0.026	0.016	0.109	−0.114
600	−0.001	−0.005	−0.023	0.013	0.187	−0.108
650	0.005	−0.007	−0.019	0.009	0.265	−0.103
700	0.010	−0.009	−0.016	0.006	−0.343	−0.097
750	0.016	−0.012	−0.012	0.002	−0.421	−0.091
800	0.021	−0.014	−0.009	−0.001	−0.500	−0.086
850	0.027	−0.016	−0.005	−0.004	−0.578	−0.08
900	0.033	−0.019	−0.002	−0.008	0.656	−0.075

(continued)

(continued)

	Attar-Schwarz (2009)	Khoury-Kassabri et al. (2004)	Mooij et al. (2011)	Van der Vegt et al. (2005)	Vieno et al. (2005)	Wei et al. (2010)
950	0.038	−0.021	0.002	−0.011	0.734	−0.069
1,000	0.044	−0.024	0.005	−0.015	0.812	−0.063
1,050	0.049	−0.026	0.009	−0.018	0.890	−0.058
1,100	0.055	−0.028	0.012	−0.021	0.968	−0.052
1,150	0.060		0.016	−0.025		−0.047
1,200	0.066		0.019	−0.028		−0.041
1,250	0.072		0.022	−0.032		−0.035
1,300			0.026	−0.035		−0.030
1,350			0.029	−0.038		−0.024
1,400			0.033	−0.042		−0.019
1,450			0.036	−0.045		−0.013
1,500			0.040	−0.049		−0.008
1,550			0.043	−0.052		−0.002
1,600			0.047	−0.055		0.004
1,650			0.050	−0.059		0.009
1,700			0.053	−0.062		0.015
1,750			0.057	−0.066		0.020
1,800			0.060	−0.069		0.026
1,850			0.064	−0.072		0.032
1,900			0.067	−0.076		0.037
1,950			0.071	−0.079		0.043
2,000			0.074	−0.083		0.048
2,050				−0.086		0.054
2,100				−0.089		0.06
2,150				−0.093		0.065
2,200				−0.096		0.071
2,250						0.076
2,300						0.082
2,350						0.087
2,400						0.093

Appendix 6c: Average Outcomes Noncognitive *Student* Scores Per School Size in Secondary Education

	Weighted average (all)	Weighted average (U.S.)	Weighted average (non-U.S.)	Weighted average (non-U.S., excluding Vieno et al.)
N (students)	106,765	42,488	64,277	59,544
School size				
300	−0.0160	0.0284	−0.0454	−0.0294
350	−0.0120	0.0266	−0.0376	−0.0257
400	−0.0079	0.0246	−0.0294	−0.0217
450	−0.0041	0.0222	−0.0214	−0.0179
500	−0.0002	0.0198	−0.0134	−0.0140
550	0.0038	0.0173	−0.0052	−0.0099
600	0.0075	0.0149	0.0026	−0.0062
650	0.0113	0.0120	0.0107	−0.0022
700	0.0152	0.0099	0.0188	0.0016
750	0.0190	0.0073	0.0267	0.0055
800	0.0230	0.0051	0.0348	0.0093
850	0.0270	0.0027	0.0430	0.0135
900	0.0308	0.0002	0.0510	0.0173
950	0.0349	−0.0016	0.0590	0.0212
1,000	0.0388	−0.0038	0.0670	0.0250
1,050	0.0429	−0.0057	0.0750	0.0288
1,100	0.0469	−0.0079	0.0832	0.0329

Appendix 7a: Predicted Noncognitive *School Mean*s Cores Per School Size in Secondary Education; American Studies

	Chen (2008)	Chen and Vazsonyi (2013)	Chen (2008)	Gottfredson and DiPietro (2011)	Haller (1992)	Kirkpatrick Johnson et al. (2001); middle schools	Kirkpatrick Johnson et al. (2001); high schools	Klein and Cornell (2010)	McNeely et al. (2002)	Payne (2012)	Rumberger and Palardy (2005)
N (schools)	712	85	212	253	558	45	64	290	127	253	912
School size											
100			0.157	−0.702	0.175		0.295	0.016	0.662	0.233	
150			0.148	−0.565	0.164	0.000	0.281	0.015	−0.518	0.188	
200	0.740	0.655	0.139	−0.467	0.154	0.000	0.267	0.015	−0.415	0.152	−0.332
250	0.665	−0.577	0.129	−0.391	0.144	0.000	0.253	0.014	−0.336	0.124	−0.302
300	−0.593	−0.501	0.120	−0.329	0.134	0.000	0.239	0.013	0.271	0.100	0.274
350	−0.524	−0.430	0.111	−0.277	0.124	0.000	0.225	0.012	0.216	0.079	0.246
400	−0.458	−0.362	0.102	−0.232	0.114	0.000	0.211	0.012	0.168	0.060	0.219
450	−0.394	0.297	0.093	−0.192	0.104	0.000	0.197	0.011	0.127	0.043	0.193
500	−0.334	0.235	0.084	−0.156	0.094	0.000	0.183	0.010	0.089	0.028	0.168
550	0.276	0.177	0.075	−0.124	0.084	0.000	0.168	0.010	0.055	0.015	0.144
600	0.221	0.123	0.066	−0.094	0.073	0.000	0.154	0.009	0.024	0.002	0.121
650	0.169	0.072	0.056	−0.067	0.063	0.000	0.140	0.008	−0.004	−0.009	0.098
700	0.120	0.024	0.047	−0.042	0.053	0.000	0.126	0.007	−0.031	−0.020	0.077
750	0.074	−0.020	0.038	−0.018	0.043	0.000	0.112	0.007	−0.055	−0.030	0.057
800	0.030	−0.061	0.029	0.003	0.033	0.000	0.098	0.006	−0.078	−0.040	0.037
850	−0.011	−0.099	0.020	0.024	0.023		0.084	0.005	−0.100	−0.049	0.019
900	−0.048	−0.133	0.011	0.043	0.013		0.070	0.004	−0.120	−0.057	0.001
950	−0.083	−0.163	0.002	0.062	0.003		0.056	0.004	−0.139	−0.066	−0.015
1,000	−0.116	−0.190	−0.007	0.079	−0.008		0.041	0.003	−0.158	−0.073	−0.031
1,050	−0.145	−0.214	−0.016	0.096	−0.018		0.027	0.002	−0.175	−0.081	−0.046
1,100	−0.171	−0.235	−0.026	0.111	−0.028		0.013	0.002	−0.192	−0.088	−0.059
1,150	−0.195	−0.251	−0.035	0.127	−0.038		−0.001	0.001	−0.208	−0.094	−0.072

(continued)

(continued)

	Chen (2008)	Chen and Vazsonyi (2013)	Chen (2008)	Gottfredson and DiPietro (2011)	Haller (1992)	Kirkpatrick Johnson et al. (2001); middle schools	Kirkpatrick Johnson et al. (2001); high schools	Klein and Cornell (2010)	McNeely et al. (2002)	Payne (2012)	Rumberger and Palardy (2005)
1,200	−0.216	−0.265	−0.044	0.141	−0.048		−0.015	0.000	−0.223	−0.101	−0.084
1,250	−0.234		−0.053	0.155	−0.058		−0.029	−0.001	−0.237	−0.107	−0.095
1,300			−0.062	0.168	−0.068		−0.043	−0.001	−0.251	−0.113	−0.105
1,350			−0.071	0.181	−0.078		−0.057	−0.002	−0.265	−0.118	−0.115
1,400			−0.080	0.193	−0.089		−0.071	−0.003	−0.278	−0.124	−0.123
1,450			−0.089	0.205	−0.099		−0.086	−0.003	−0.290	−0.129	−0.130
1,500			−0.099	0.217	−0.109		−0.100	−0.004	−0.302	−0.134	−0.136
1,550			−0.108	0.228	−0.119		−0.114	−0.005	−0.314	−0.139	−0.142
1,600			−0.117	0.239	−0.129		−0.128	−0.006	−0.325	−0.144	−0.146
1,650			−0.126	0.249	−0.139		−0.142	−0.006	−0.336	−0.149	−0.150
1,700			−0.135	0.259	−0.149		−0.156	−0.007	−0.347	−0.154	−0.152
1,750			−0.144	0.269	−0.159		−0.170	−0.008	−0.357		−0.154
1,800			−0.153		−0.169		−0.184	−0.009	−0.367		−0.155
1,850			−0.162		−0.180		−0.198	−0.009	−0.377		−0.155
1,900			−0.172		−0.190		−0.213	−0.010	−0.386		−0.153
1,950			−0.181		−0.200		−0.227	−0.011	−0.396		−0.151
2,000					−0.210		−0.241	−0.011	−0.405		−0.148
2,050					−0.220		−0.255	−0.012	−0.413		−0.144
2,100					−0.230		−0.269	−0.013	−0.422		−0.140
2,150					−0.240		−0.283	−0.014	−0.430		−0.134
2,200					−0.250		−0.297	−0.014	−0.438		−0.127
2,250					−0.261		−0.311	−0.015			
2,300					−0.271		−0.325	−0.016			
2,350					−0.281		−0.340	−0.017			
2,400					−0.291		−0.354	−0.017			

Appendix 7b: Predicted Noncognitive *School Mean* Scores Per School Size in Secondary Education; Studies Outside the U.S

	Attar-Schwarz (2009)	Khoury-Kassabri et al. (2004)	Mooij et al. (2011)	Silins and Mulford (2004)	Vieno et al. (2005)	Wei et al. (2010)
N (schools)	327	162	104	96	134	12
School size						
100	−0.220	0.060	−0.272	−0.437	−2.967	−0.423
150	−0.198	0.053	−0.255	−0.396	−2.577	−0.408
200	−0.177	0.045	−0.239	−0.355	−2.187	−0.394
250	−0.155	0.038	−0.222	−0.314	−1.796	−0.379
300	−0.133	0.031	−0.206	0.272	−1.406	−0.365
350	−0.112	0.023	−0.189	0.231	−1.016	−0.351
400	−0.090	0.016	−0.173	0.190	−0.625	−0.336
450	−0.068	0.008	−0.156	0.149	−0.235	−0.322
500	−0.047	0.001	−0.140	0.108	0.155	−0.307
550	−0.025	−0.007	−0.124	0.067	−0.546	−0.293
600	−0.003	−0.014	−0.107	0.026	0.936	−0.279
650	0.018	−0.021	−0.091	−0.015	1.327	−0.264
700	0.040	−0.029	−0.074	−0.056	1.717	−0.250
750	0.062	−0.036	−0.058	−0.097	2.107	−0.235
800	0.083	−0.044	−0.041	−0.138	20498	−0.221
850	0.105	−0.051	−0.025	−0.179	2.888	−0.207
900	0.127	−0.059	−0.009	−0.220	3.278	−0.192
950	0.148	−0.066	0.008	−0.261	3.669	−0.178
1,000	0.170	−0.073	0.024	−0.302	4.059	−0.163
1,050	0.192	−0.081	0.041	−0.343	4.449	−0.149
1,100	0.213	−0.088	0.057	−0.384	4.84	−0.135
1,150	0.235		0.074	−0.425		−0.120
1,200	0.257		0.090	−0.466		−0.106
1,250	0.278		0.107			−0.091
1,300			0.123			−0.077
1,350			0.139			−0.063
1,400			0.156			−0.048
1,450			0.172			−0.034
1,500			0.189			−0.019
1,550			0.205			−0.005
1,600			0.222			0.009
1,650			0.238			0.024
1,700			0.254			0.038
1,750			0.271			0.053
1,800			0.287			0.067
1,850			−0.304			0.081

(continued)

(continued)

	Attar-Schwarz (2009)	Khoury-Kassabri et al. (2004)	Mooij et al. (2011)	Silins and Mulford (2004)	Vieno et al. (2005)	Wei et al. (2010)
1,900			−0.320			0.096
1,950			−0.337			0.110
2,000			−0.353			0.125
2,050						0.139
2,100						0.153
2,150						0.168
2,200						0.182
2,250						0.197
2,300						0.211
2,350						0.225
2,400						0.240

Appendix 7c: Average Outcomes Noncognitive *School Mean* Scores Per School Size in Secondary Education

	Weighted average (all)	Weighted average (U.S.)	Weighted average (non-U.S.)	Weighted average (non-U.S. excluding Vieno)
N	(schools)	4346	3511	835
701				
School size				
200	0.151	0.283	−0.107	−0.066
250	0.142	0.257	−0.098	−0.060
300	0.134	0.231	−0.089	−0.054
350	0.127	0.206	−0.080	−0.049
400	0.120	0.181	−0.070	−0.044
450	0.114	0.158	−0.061	−0.038
500	0.109	0.135	−0.052	−0.033
550	0.104	0.114	−0.042	−0.028
600	0.100	0.093	−0.032	−0.022
650	0.097	0.073	−0.022	−0.017
700	0.094	0.054	−0.012	−0.011
750	0.093	0.036	−0.002	−0.005
800	0.092	0.018	0.008	0.000
850	0.091	0.002	0.019	0.005
900	0.092	−0.014	0.030	0.011
950	0.093	−0.028	0.040	0.016
1,000	0.095	−0.042	0.051	0.022
1,050	0.097	−0.055	0.062	0.027
1,100	0.101	−0.067	0.073	0.033

References

Included in the Quantitative Summary of Research Findings

Archibald, S. (2006). Narrowing in on educational resources that do affect student achievement. *Peabody Journal of Education*, *81*(4), 23–42. http://dx.doi.org/10.1207/s15327930pje8104_2

Attar-Schwartz, S. (2009). Peer sexual harassment victimization at school: The Roles of student characteristics, cultural affiliation, and school factors. *American Journal of Orthopsychiatry, 79*(3), 407–420. doi:10.1037/a0016553

Bonnet, M., Gooss, F. A., Willemen, A. M., & Schuengel, C. (2009). Peer victimization in dutch school classes of four- to five-year-olds: Contributing factors at the school level. *The Elementary School Journal, 110*(2), 163–177.

Bowen, G. L., Bowen, N. K., & Richman, J. M. (2000). School size and middle school students' perceptions of the school environment. *Social Work in Education, 22*(2), 69–82.

Carolan, B. V. (2012). An examination of the relationship among high school size, social capital, and adolescents' mathematics achievement. *Journal of Research on Adolescence, 22*(3), 583–595. doi:10.1111/j.1532-7795.2012.00779.x

Chen, G. (2008). Communities, students, schools and school crime—A confirmatory study of crime in US high schools. *Urban Education, 43*(3), 301–318. doi:10.1177/0042085907311791

Chen, P., & Vazsonyi, A. T. (2013). Future orientation, school contexts, and problem behaviors: A multilevel study. *Journal of Youth and Adolescence, 42*(1), 67–81. doi:10.1007/s10964-012-9785-4

Dee, T. S., Ha, W., & Jacob, B. A. (2007). *The effects of school size on parental involvement and social capital: Evidence from the ELS:2002*. Brookings Papers on Education Policy, pp. 77–97.

Durán-Narucki, V. (2008). School building condition, school attendance, and academic achievement in New York City public schools: A mediation model. *Journal of Environmental Psychology, 28*(3), 278–286. doi:10.1016/j.jenvp.2008.02.008

Fernandez, K. E. (2011). Evaluating school improvement plans and their affect on academic performance. *Educational Policy, 25*(2), 338–367. doi:10.1177/0895904809351693

Gottfredson, D. C., & DiPietro, S. M. (2011). School size, social capital, and student victimization. *Sociology of Education, 84*(1), 69–89. doi:10.1177/0038040710392718

Haller, E. J. (1992). High-school size and student indiscipline: Another aspect of the school consolidation issue. *Educational Evaluation and Policy Analysis, 14*(2), 145–156.

Holas, I., & Huston, A. C. (2012). Are middle schools harmful? The role of transition timing, classroom quality and school characteristics. *Journal of Youth and Adolescence, 41*(3), 333–345. doi:10.1007/s10964-011-9732-9

Khoury-Kassabri, M., Benbenishty, R., Astor, R. A., & Zeira, A. (2004). The contributions of community, family, and school variables to student victimization. *American Journal of Community Psychology, 34*(3–4), 187–204.

Kirkpatrick Johnson, M., Crosnoe, R., & Elder, G. H, Jr. (2001). Students' attachment and academic engagement: the role of race and ethnicity. *Sociology of Education, 74*(4), 318–340.

Klein, J., & Cornell, D. (2010). Is the link between large high schools and student victimization an illusion? *Journal of Educational Psychology, 102*(4), 933–946. doi:10.1037/a0019896

Koth, C. W., Bradshaw, C. P., & Leaf, P. J. (2008). A multilevel study of predictors of student perceptions of school climate: The effect of classroom-level factors. *Journal of Educational Psychology, 100*(1), 96–104. doi:10.1037/0022-0663.100.1.96

Lee, V. E., & Loeb, S. (2000). School size in Chicago elementary schools: Effects on teachers' attitudes and students' achievement. *American Educational Research Journal, 37*(1), 3–31.

Lee, V. E., & Smith, J. B. (1997). High school size: Which works best and for whom? *Educational Evaluation and Policy Analysis, 19*(3), 205–227.

Luyten, H. (1994). School size effects on achievement in secondary education: Evidence from the Netherlands, Sweden, and the USA. *School Effectiveness and School Improvement, 5*(1), 75–99. doi:10.1080/0924345940050105

Ma, X., & McIntyre, L. J. (2005). Exploring differential effects of mathematics courses on mathematics achievement. *Canadian Journal of Education/Revue Canadienne de l'éducation, 28*(4), 827–852.

Maerten-Rivera, J., Myers, N., Lee, O., & Penfield, R. (2010). Student and school predictors of high-stakes assessment in science. *Science Education, 94*(6), 937–962. doi:10.1002/sce.20408

McNeely, C. A., Nonnemaker, J. M., & Blum, R. W. (2002). Promoting school connectedness: Evidence from the National Longitudinal Study of Adolescent Health. *Journal of School Health, 72*(4), 138–146.

Mooij, T., Smeets, E., & de Wit, W. (2011). Multi-level aspects of social cohesion of secondary schools and pupils' feelings of safety. *British Journal of Educational Psychology, 81*(3), 369–390. doi:10.1348/000709910X526614

Payne, A. A. (2012). Communal school organization effects on school disorder: Interactions with school structure. *Deviant Behavior, 33*(7), 507–524. doi:10.1080/01639625.2011.636686

Ready, D. D., & Lee, V. E. (2007). *Optimal context size in elementary schools: Disentangling the effects of class size and school size and school size*. Brookings Papers on Education Policy, pp. 99–135.

Rumberger, R. W., & Palardy, G. J. (2005). Test scores, dropout rates, and transfer rates as alternative indicators of high school performance. *American Educational Research Journal, 42*(1), 3–42.

Silins, H., & Mulford, B. (2004). Schools as learning organisations—Effects on teacher leadership and student outcomes. *School Effectiveness and School Improvement*, *15*(3–4), 443–466. http://dx.doi.org/10.1080/09243450512331383272

van der Vegt, A. L., den Blanken, M., & Hoogeveen, K. (2005). *Nationale scholierenmonitor: meting voorjaar 2005*. Utrecht: Sardes.

Vieno, A., Perkins, D. D., Smith, T. M., & Santinello, M. (2005). Democratic school climate and sense of community in school: A multilevel analysis. *American Journal of Community Psychology, 36*(3–4), 327–341. doi:10.1007/s10464-005-8629-8

Wei, H. S., Williams, J. H., Chen, J. K., & Chang, H. Y. (2010). The effects of individual characteristics, teacher practice, and school organizational factors on students' bullying: A multilevel analysis of public middle schools. *Children and Youth Services Review, 32*(1), 137–143. doi:10.1016/j.childyouth.2009.08.004

Wyse, A. E., Keesler, V., & Schneider, B. (2008). Assessing the effects of small school size on mathematics achievement: A propensity score-matching approach. *Teachers College Record, 110*(9), 1879–1900.

Chapter 5
Summary and Discussion

Hans Luyten

5.1 Research Findings

The prior chapters of this review have focused on the relation between school size in primary and secondary education and three types of outcomes, namely academic achievement, noncognitive outcomes and per pupil expenditure. The second chapter comprised a narrative discussion of prior reviews and paid special attention to aspects of school size in an international context. Previous reviews have not yet produced a consistent picture on the relation between school size and educational output (either cognitive or noncognitive), but given the large number of nonsignificant findings in studies on school size effects it does not seem likely that the relation between school size and academic achievement is very strong. Several reviews conclude that the effect of school size matters mostly for disadvantaged students. As far as estimates of optimum school sizes are specified they tend to cover a broad range. The evidence for effects of school size seems somewhat stronger for noncognitive outcomes than for academic achievement. The research literature on the relation between cost and school size appears to be fairly consistent and indicates lower expenditure per pupil in large schools.

In general, it seems that school size effects may depend quite strongly on modifying conditions. Background of the students and type of outcomes (cognitive vs. noncognitive) have already been mentioned as relevant factors in this respect and the level of education (primary vs. secondary) may also be a highly relevant factor. Attempts to identify indirect effects of school size (e.g., via a more personalized climate or a more focused curriculum) have hardly been made in this line of research.

The evidence provided by internationally comparative datasets (especially the PISA surveys) suggests a weak relation between academic achievement and school size. Moreover, for most countries the relation appears to be positive rather than negative. This implies relatively high scores in large schools and poor results in smaller schools. The findings from international surveys also indicate that in the Netherlands the average size of primary schools clearly falls below the cross-national average among OECD-countries, whereas Dutch secondary schools are relatively large in comparison to other OECD countries.

H. Luyten et al., *School Size Effects Revisited*, SpringerBriefs in Education,
DOI: 10.1007/978-3-319-06814-5_5, © The Author(s) 2014

The third chapter provides a summary of the research literature based on vote counts combined with short abstracts of research reports. The review in this chapter is based on 84 studies, which relate to 107 samples producing 277 effect size estimates. With regard to academic achievement it is found that the bulk of the reported school size effects fail to reach statistical significance. This is in line with the (tentative) conclusion in the first chapter that, even if school size does affect student achievement, the effect is not likely to be very strong. Significantly negative, positive, and curvilinear effects have been found. Negative effects have been most frequently reported.

With regard to noncognitive outcomes the vote count review distinguishes between the following six types:

- attitudes of students and teachers toward school (social cohesion)
- participation of students, teachers, or parents
- school safety (disciplinary climate, bullying, problem behavior, and violence)
- student absence and dropout
- other student outcomes (attitudes toward self or learning; engagement)
- school organization, teaching, and learning

The overall picture that emerges from the vote count analysis on noncognitive outcomes is that nearly half of all the school size effects reported in the reviewed studies are statistically significant and negative. The remaining effects are mostly nonsignificant, although positive and curvilinear effects have been reported as well. All in all the available evidence indicates that small schools tend to show positive associations with noncognitive outcomes. However, the fact that in many studies findings have been reported that deviate from this general picture, suggest a fairly modest effect of school size on noncognitive outcomes.

With respect to attitudes 14 studies (including 14 samples) were reviewed. These studies reported findings with regard to 24 effects, of which 17 were found to be statistically significant and negative. This clearly suggests a positive impact of small schools on student and teacher attitudes. Ten studies that focused on participation were reviewed. Also in this case a large majority of the effects turned out to be negative and significant. Regarding school safety 24 studies were reviewed and these reported details on 54 effects. A large number of the effects (22) were statistically not significant and a nearly equally large number (21) were significantly negative. The remaining effects were either curvilinear or significantly positive. For safety the effects therefore tend to be negative (i.e., more safety in smaller schools), but the evidence is less convincing than for attitudes and participation. It was also evident from the vote count analysis that negative effects of school size were relatively often reported in the United States. Regarding absence and dropout 16 studies were reviewed reporting on 28 effects. Also in this case a large number (13) of the reported effects were found to be nonsignificant. A slightly smaller number (11) of effects appeared to be significantly negatively. The remaining effects were either positive or curvilinear. Also in this respect the available evidence shows more support in favor of small schools, but deviating findings are quite frequently reported as well. Six studies were reviewed that relate

to student engagement or attitudes toward self or learning. Most findings appeared not be statistically significant. No evidence for a positive school size effect was found, but a small number of negative and curvilinear effects were reported. With respect to school organization, teaching, and learning four studies were reviewed. Of the eighteen effects reported eleven were found not to be significant. Six of the remaining effects appeared to be negative and one curvilinear.

The research evidence is quite clear-cut with regard to per pupil expenditure and school size in primary and secondary education. All five studies that were reviewed revealed a similar pattern. Costs per pupils tend to decline as schools get larger. This is especially the case for relatively small schools. The prospects of cost reduction through increasing school size are rather modest for schools of average size or larger. It is important to note that this conclusion is based on studies that control for educational output (e.g., academic achievement or graduation rates). Only one study (Stiefel et al. 2000) suggests that the relation between school size and per pupil expenditure may not hold when controlling for student population characteristics. Although the available evidence reveals a consistent pattern, it should be noted that the actual expenditure per pupils strongly depends on local regulations with regard to school funding and teacher salaries. Staff salaries make up the bulk of educational expenditure and salary levels for teachers are often primarily based on work experience and qualification levels. Besides, specific regulations to support small schools (e.g., specific funding for religious schools or schools in rural communities) may also influence the relation between school size and per pupil expenditure in certain settings.

In the fourth chapter the effect of school size on student achievement and noncognitive outcomes has been summarized in a quantitative manner. Per school size the average standardized outcome across a number of studies has been reported. The studies included in this summary form a subset of the studies covered in the second chapter. Only if the information needed to calculate the "predicted" outcomes at a given school size was reported, could a study be included. Furthermore, requirements were made regarding the control variables included in the analyses. Studies on academic achievement were only included if the outcomes were controlled for prior achievement or if they related to learning gains. For studies with regard to noncognitive outcomes controlling for socioeconomic background was deemed sufficient for inclusion in the summary. The discussion of the conclusions from this summary will be largely confined to the effect of school size on outcomes that are standardized by dividing them through the standard deviation in individual level outcomes (in the third chapter referred to as student outcomes). As explained earlier we consider this standardization the most appropriate. The alternative (standardization by means of the standard deviation in school means) is bound to produce larger effects, but fails to take into account that an effect of school size is by definition largely limited to the school level and can only affect the variation between students within schools in interaction with student level variables (e.g., if the effect of school size is stronger for socioeconomically disadvantaged students). Separate findings are presented for primary and secondary education. With respect to noncognitive outcomes in

secondary education the findings are reported separately for American studies and for studies conducted in other countries. It should be noted that the findings reported for academic achievement (both primary and secondary education) and noncognitive outcomes in primary education are predominantly based on findings from American research.

With regard to academic achievement in primary education the findings suggest a negative, but very weak effect of school size. The findings indicate a difference that is smaller than one tenth of a standard deviation between schools with 200 students and schools with 800 students. For achievement scores in secondary education the findings suggest a curvilinear pattern. The highest scores are found in schools with at least 1,200 students but less than 1,600. If schools grow larger achievement scores tend to decrease, but the lowest scores occur in very small schools (i.e., less than 700 students). The difference between the lowest scoring schools (400 students) and the highest scoring (1350–1500 students) is just over one tenth of a standard deviation. The effect of school size cannot easily be expressed in a more current metric (e.g., Cohen's d or a correlation coefficient), because its association with outcomes does not always follow a linear pattern. With regard to academic achievement in secondary education, it seems plausible that the difference between the highest and lowest scoring schools is less than one tenth of a standard deviation. This is commonly considered a (very) small effect (i.e., Cohen's $d < 0.10$[1]). Our review implies that the other school size effects covered in this review (i.e., on outcomes in primary education and noncognitive outcomes in secondary education) are hardly any stronger. The (lack of) practical significance of such an effect may be illustrated by the following example. Assume that the effect relates to a test with a mean score of 50 and a standard deviation of 10. This would imply that 95 % of the scores falls within the 30–70 range (assuming a standard normal distribution of the test scores). The findings from our review imply that it would require massive changes in school size (by at about one thousand students) to arrive at a change of more than two points.

With regard to noncognitive outcomes in primary education the findings suggest a negative, but fairly weak effect. The effect of school size in primary education appears to be somewhat stronger than what was found for academic achievement. The difference between a school with 200 versus a school with 600 students amounts to 0.129 standard deviation. With regard to academic achievement this difference would be 0.074 standard deviation (also see Table 4.4 and Appendix 4 of Chap. 4). In terms of Cohen's d these differences also represent (very) small effects ($d = 0.13$ and $d = 0.07$ respectively).

When it comes to the relation between school size and noncognitive outcomes in secondary education, clearly opposite trends are apparent in American studies versus studies from other countries. The general trend across all studies is slightly in favor of large schools. The difference between small secondary schools

[1] As a rule of thumb an effect size equal to 0.20 is considered small, 0.50 is considered medium and 0.80 is considered large.

(300 students) and large ones (1,100 students) amounts to 0.06 standard deviation (see Appendix 6c of Chap. 4). When the summary is based exclusively on American studies, the trend is reversed. Small schools show more favorable scores, although the difference between small and large American schools turns out to be very modest (0.04 standard deviation). The effect of school size in non-US studies is somewhat stronger and reversed (showing more positive scores in large schools). The differences between schools with 300 students versus schools with 1,100 students equals 0.06 standard deviation.[2]

5.2 Research Questions Revisited

In the introduction the following four research questions were formulated:

(1) What is the impact of school size on cognitive learning outcomes, noncognitive outcomes, and the social distribution of learning outcomes?
(2) What is the "state of the art" of the empirical research on economies of size?
(3) What is the direct and indirect impact of school size, conditioned by other school context variables on student performance? (where indirect effects are perceived as influencing through intermediate school and instruction characteristics)
(4) What is the specific position of the Netherlands in international perspective?

With regard to the first question it can be concluded from the reported findings that the impact of school size on cognitive and noncognitive outcomes is weak. The strongest effects have been found for social cohesion and participation. Some American studies report that school size effects are relatively strong for socio-economically disadvantaged students.

With regard to economies of size (question 2) the studies reviewed show a consistent pattern of decreasing expenditures per pupil as school size increases. The financial benefits are not outweighed by decreasing educational output (such as academic achievement or graduation rates). Reductions in costs are relatively modest if a school increases in size from average to above average, but they are more substantial if the size of a (very) small school increases.

With regard to question 3 the conclusion is that given the variation in research findings many factors that are not yet well understood might interfere with the impact of school size. Many of these interfering factors (preconditions or intermediary variables) are not yet identified, although a few can be specified. Clearly the type of education (primary/secondary) is an important precondition. Chapter 4 shows that, as far as optimal school sizes can be discerned, they obviously differ between primary and secondary education. Some American studies suggest that

[2] This difference excludes the findings from the study by Vieno et al. (2005), who reported a very strong but non-significant effect of school size.

the socioeconomic backgrounds of student populations may affect the strength of a school size effect. Our review of the research literature also shows that the evidence of school size effects is relatively strong for some noncognitive outcomes (social cohesion and participation) and less so for other outcomes (e.g., academic achievement). It is also clear that the national context plays an important role. International comparisons suggest that several of the negative school size effects that have been reported in American studies may actually work in the opposite direction in other countries. With regard to intermediary variables, our present knowledge is still limited as structural equation modeling (SEM) has thus far only been applied in a handful of studies. However, these studies provide some support for the idea that a negative effect of school size on academic achievement is mediated by social climate.

With regard to the fourth question it can be concluded that the international comparisons reported in Chap. 1 indicate that in Dutch primary education schools are relatively small in comparison to other OECD-countries. This may be conceived as an opportunity to cut costs by increasing the size of primary schools. Most gains could be realized by merging the smallest schools. A complicating factor in this respect is that most small schools are situated in small and rural communities. Closing down schools in such communities has serious implications for its inhabitants. The disappearance of schools may turn a rural community into a mere collection of houses situated close to each other. Such consequences should be taken into account before making decisions that are purely based of financial considerations. Moreover, it must be noted that the findings reported in the fourth chapter indicate that in primary education the most beneficial results are obtained in small schools. The findings indicate that the best results are found in schools with no more than 200 students. As the average school size in Dutch primary education is somewhat higher (but still below 300 students), there is little reason for increasing school size.

With regard to secondary education our findings indicate that the best results are found in relatively large schools (over 1,200 students), but that too large may be detrimental. Although the average school size in Dutch education is relatively large, it is still less than the size that appears to produce the best results across the studies reviewed for the present report.

5.3 Policy Implications

In our view the research base on school size effects hardly gives any reasons for policy measures. The outcomes to be expected from decreasing school size on either cognitive or noncognitive outcomes will first of all be rather modest. More importantly, given the substantial variation in research outcomes, the results to be expected from a policy effort aimed at school size decreases are highly uncertain. In view of the large variation in reported outcomes across studies, one has to conclude that both modifying preconditions or intermediating factors play an

important role in the path from changes in school size to outcomes. However, in this respect our knowledge base is still very weak. The present review has shown that school size effects differ between primary and secondary education and also across educational systems. Most of the evidence on negative effects of large school size on both cognitive and noncognitive outcomes derives from studies conducted in the US Much of the research that relates to non-US contexts even suggests an effect in the opposite direction. Not only are we dealing with a lack of knowledge on the preconditions that may affect the impact of school size, but also the causal mechanisms (the "how and why") that can account for the presumed impact of school size, about which very little is known. With regard to academic achievement a more personalized atmosphere as well as scope and focus of the school curriculum are sometimes mentioned as mediating variables. Research methods that focus explicitly on modeling such indirect effects (such as structural equation modeling) have thus far been applied in no more than a handful of studies on school size. Only three studies that were reviewed for the present report provide information on the role of mediating variables between school size and outcomes that is based on empirical research (Chen and Weikart 2008; Coladarci and Cobb 1996; Silins and Mulford 2004). Although these studies provide some support for the idea that school size may affect more distal outcomes via a direct impact on cohesion and participation, our understanding of how and why school size produce certain effects is still rather incomplete. If school size is considered as an instrument for enhancing academic achievement and noncognitive outcomes, better knowledge on how, why, and under what circumstances it may work would be welcome. Given our present knowledge, the results of changes in school size are difficult to predict.

The uncertainty about what changes in school size might bring about apply most strongly to academic outcomes. With respect to some of the noncognitive outcomes (especially participation and attitudes that affect social cohesion) the findings from empirical research are more consistent. A considerable number of studies have come up with insignificant findings, but the large majority of the significant findings point to negative effects (small schools do better on these outcomes).

Research on the relation between school size and per student expenditure was found to be highly consistent across studies. These findings imply that increasing school size is likely to decrease the expenditure per pupil. It should be noted that the sharpest decreases in cost reduction can be expected if the size of schools smaller than average increases. The advantages of school size increases grow less and less as schools get larger. It is important to note that studies on the relation between school size and cost per pupil typically include measures of educational output (e.g., achievement scores or graduation rates) as a covariate in their analyses. As such the reported effects express financial gains given a constant level of educational output. This implies that the risk of diseconomies of scale (i.e., financial gains do not weigh up to loss in educational output) must be considered quite small. Still, it needs to be acknowledged that the actual costs per students in a given context are strongly determined by the level of teacher salaries (which

mainly depend on work experience and qualifications) and the pupil teacher ratio. In many educational systems the pupil teacher ratio depends strongly on government regulations that specify the amount of funding a school receives for hiring teaching staff. The number of students enrolled is usually the main factor that determines how many teachers a school can hire. Also specific regulations for founding and closing schools may be a relevant factor. In the Netherlands, permission to start a school does not only depend on the number of pupils that a school will enroll, but also the religious denomination and educational philosophy that serve as criteria.

In our view, the impact of school size on educational outcomes (cognitive or non-cognitive) should not be overrated. Attempts to improve these aspects should focus on factors with a close link with achievement outcomes (quality of instruction; introducing and maintaining clear rules about behavior in school). Manipulating school size (downsizing or otherwise) should not be the next hobby for managers and administrators. It may be an attractive instrument for managers as they can easily show what they have "accomplished". However, the accomplishment that should be the focus of their attention is better instruction and improved personal relations, which are much more difficult to demonstrate by means of quantitative indicators.

5.4 Suggestions for Future Research on School Size

The present review indicates that effects of school size on cognitive and non-cognitive outcomes are modest in general. Perhaps even more important is the lack of certainty regarding the impact of changes in school size. The variation in findings across numerous studies suggests that several unknown confounding variables affect the course from school size to educational outcomes. Such variables may entail preconditions that either facilitate or impede the effect. They may also relate to intermediating variables. The longer the causal chain from school size to outcomes is and the larger the number of intermediating variables are, the more uncertain the eventual effect of school size becomes. The effect of school size on relevant intermediating variables (e.g., participation, social cohesion) will be uncertain to some extent and the same goes for the effect of the intermediating variables on the outcomes. Increasing our knowledge of such preconditions and intermediating variables is likely to shed more light on the why and how of school size effects. At present there still is a remarkable gap in our knowledge base. Reduction (or expansion) of school size as a policy instrument seems hazardous, if it is unclear what processes are actually set in motion by changes in school size. Therefore, future research should not primarily aim at the eventual outcomes of school size, but rather try to clarify the role of preconditions and mediating variables. Structural equation modeling (SEM) might serve a promising methodology in this respect. Testing causal models of school size effects may help to disentangle the relations among several variables affected by (and affecting) school

size. In addition, more studies on actual changes in school size over time would be welcome contributions. Thus far, most of the research has been based on comparing schools of different sizes. Studies on changes over time provide stronger evidence for causal relations. Finally, more research outside the US is badly needed. The debate on school size is largely inspired by research findings from American studies, but our findings suggest that the US may not be representative at all for most other countries.

References

Chen, G., & Weikart, L. A. (2008). Student background, school climate, school disorder, and student achievement: An empirical study of New York city's middle schools. *Journal of School Violence, 7*(4), 3–20. http://dx.doi.org/10.1080/15388220801973813

Coladarci, T., & Cobb, C. D. (1996). Extracurricular participation, school size, and achievement and self-esteem among high school students: A national look. *Journal of Research in Rural Education, 12*(2), 92–103.

Silins, H., & Mulford, B. (2004). Schools as learning organisations—Effects on teacher leadership and student outcomes. *School Effectiveness and School Improvement, 15*(3–4), 443–466. http://dx.doi.org/10.1080/09243450512331383272

Stiefel, L., Berne, R., Iatarola, P., & Fruchter, N. (2000). High school size: Effects on budgets and performance in New York City. *Educational Evaluation and Policy Analysis, 22*(1), 27–39.

Vieno, A., Perkins, D. D., Smith, T. M., & Santinello, M. (2005). Democratic school climate and sense of community in school: A multilevel analysis. *American Journal of Community Psychology, 36*(3–4), 327–341. doi:10.1007/s10464-005-8629-8