THE CAMPBELL COLLABORATION

Title Registration for Systematic Review: Small Class Sizes for Improving Student Achievement in Primary and Secondary Schools

Trine Filges, Christoffer Scavenius Sonne-Schmidt, Tine Nielsen, & Ann Marie Klint Jørgensen

Submitted to the Coordinating Group of:

☐	Crime and Justice
☒	Education
	☐ Disability
☐	International Development
	☐ Nutrition
☐	Social Welfare
☐	Other:

Plans to co-register:

☒	No		
☐	Yes	☐ Cochrane	☐ Other
☐	Maybe		

Date Submitted: 6 March 2012

Date Revision Submitted: 4 July 2012

Approval Date: 9 October 2012

Publication Date: 01 November 2012

Small Class Sizes for Improving Student Achievement in Primary and Secondary Schools

BACKGROUND

The Problem

Class size is one of the most researched educational interventions in social science, yet there is no clear consensus on the effectiveness of small class sizes for improving student achievement. While one strand of class size research points to small and insignificant effects, another points to positive and significant effects.

The early meta-analysis by Glass and Smith (1979) analysed the outcomes of 77 studies including 725 comparisons between smaller and larger class sizes on student achievement. They concluded that a class size reduction had a positive effect on student achievement. Hedges and Stock (1983) reanalysed Glass and Smith's data using different statistical methods, but found very little difference in the average effect sizes across the two analysis methods.

However, the updated literature reviews by Hanushek (Hanushek, 1989; 1999; 2003) cast doubt on these findings. His reviews looked at 276 estimates of pupil-teacher ratios as a proxy for class size, and most of these estimates pointed to insignificant effects. Based on a vote counting method, Hanushek concluded that "there is no strong or consistent relationship between school resources and student performance" (Hanushek, 1987, p. 47). Krueger (2003), however, points out that Hanushek relies too much on a few studies, which reported many estimates from even smaller subsamples of the same dataset. Many of the 276 estimates were from the same dataset but estimated on several smaller subsamples, and these many small sample estimates are more likely to be insignificant. The vote counting method used in Hanushek's original literature review (Hanushek, 1989) is also criticised by Hedges, Laine, and Greenwald (1994), which offers a reanalysis of the data from Hanushek's reviews using more sophisticated synthesis methods. Hedges et al. (1994) used a combined significance test.[1] They tested two null hypotheses: 1) no positive relation between the resource and output and 2) no negative relation between the resource and output. The tests determine if the data are consistent with the null hypothesis in all studies or false in at least some of the studies. Further, Hedges et al. (1994) reported the median standardized regression coefficient.[2] The conclusion is that "it shows systematic positive relations between resource inputs and school outcomes" (Hedges et al., 1994, p. 5). Hence, dependent

---

[1] The inverse chi-square (Fisher) method (Hedges & Olkin, 1985)
[2] The standardized regression coefficient measures the number of standard deviations of change in output that would be associated with a one standard deviation change in input.

upon which synthesis method[3] is considered appropriate, conclusions based on the same evidence are quite different.

The divergent conclusions of the above-mentioned reviews are further based on non-experimental evidence, combining measurements from primary studies that have different specifications and assumptions. According to Grissmer (1999), the different specifications and assumptions, as well as the appropriateness of the specifications and assumptions, account for the inconsistency of the results of the primary studies.

The Tennessee STAR experiment provides rare evidence of the effect of class size from a randomized controlled trial (RCT). The STAR experiment was implemented in Tennessee in the 1980s, assigning kindergarten children to either normal sized classes (ca. 22 students) or small classes (ca. 15 students). The study ran for four years, until the assigned children reached third grade, but not even based on this kind of evidence do researchers agree about the conclusion.

According to Finn and Achilles (1990), Nye, Hedges, and Konstantopoulos (1999) and Krueger (1999), STAR results show that class size reduction increased student achievement. However, Hanushek (1999; 2003) questions these results because of attrition from the project, crossover between treatments, and selective test taking, which may have violated the initial randomization.

While the class size debate on what can be concluded based on the same evidence is acceptable and meaningful in the research community, it is probably of less help in guiding decision-makers and practitioners. If research is to inform practice, there must be an attempt to reach some agreement about what the research does and does not tell us about the effectiveness of interventions as well as what conclusions can be reasonably drawn from research. The researchers must reach a better understanding of questions such as: for whom does class size reduction have an effect? When does class size reduction have an effect? How small does a class have to be in order to be advantageous?

The purpose of this review is to systematically uncover relevant studies in the literature that measure the effects of class size on academic achievement. We will synthesize the effects in a transparent manner and, where possible, we will investigate the extent to which the effects differ among different groups of students such as high/low performers, high/low income families, or members of minority/non-minority groups, and whether timing, intensity, and duration have an impact on the magnitude of the effect.

---

[3] The vote counting method did not necessarily lead to a different conclusion. It depends upon the inference procedure associated with the method (i.e., the category with the most "votes" represents the true state). The analysis in Hanushek (1989) shows that 24 (3) percent of the coefficients on expenditure were positive (negative) and significant and 46 (24) percent were positive (negative) and non-significant, implying that the typical relation is positive. Note that none of the methods used in either Hanushek (1989) or Hedges et al. (1994) combines magnitude of effect size and significance.

As it is costly to reduce class size, it is important to consider the types of students who might benefit most from smaller class sizes and to consider the timing, intensity, and duration of class size reduction as well. Smaller class size has been shown to benefit at-risk students, minority students, students living in poverty, and students who are educationally disadvantaged the most (Finn, 2002).

In the case of timing of class size reduction, the question is: when does class size reduction have the largest effect? According to Bascia and Fredua-Kwarteng (2008), researchers agree that class size reduction is most effective in the primary grades. For intensity, the question is how small does a class have to be in order to optimize the advantage? For example, large gains are attainable when class size is below 20 students (Biddle & Berliner, 2002; Finn, 2002). Finally, researchers (Biddle & Berliner, 2002; Finn, 2002; Grissmer, 1999; Nye, et al., 1999) agree that the length of the intervention is linked with sustainability of benefits.

INTERVENTION

It is important to distinguish between what is meant by class size and a student-teacher ratio. Class size is the number of students a teacher instructs in a classroom at a point in time. A student-teacher ratio is different because it is a school's total student enrollment divided by the number of its full time teachers.

In this review, the intervention is class size. Studies only considering student-teacher ratio will not be eligible. Outcomes of students in small classes will be compared with outcomes of students in large classes.

POPULATION

The review will include children in grades kindergarten to 12 (or the equivalent in European countries) in general education. Studies that meet inclusion criteria will be accepted from all countries. In this review, we exclude children in home-school, in pre-school programs, and in special education.

OUTCOMES

The primary focus is on measures of academic achievement. Academic achievement outcomes include reading and mathematics. The outcomes can be measured by standardized test scores.

Secondary outcomes related to student engagement will be included to the extent they are reported in the studies.

Time points for measures considered will be:

- Students currently in small classes

- 0 to 1 year follow up

- 1 to 2 year follow up

- More than 2 year follow up

OBJECTIVES

The main objective of this review is to evaluate current evidence on the effects of class size on student academic achievement. The aim of this review is to uncover and synthesize relevant studies in the literature that measure the effects of class size on academic achievement. Where possible, we will also investigate the extent to which the effects differ among different groups of students such as high/low performers, high/low income families, or members of minority/non-minority groups, and whether timing, intensity, and duration have an impact on the magnitude of the effect.

METHODOLOGY

The proposed project will follow standard procedures for conducting systematic reviews using meta-analysis techniques.

We will include study designs that use a well-defined control group. The study designs we will include in the review are:

- Controlled trials (all parts of the study are prospective, i.e. identification of participants, assessment of baseline, allocation to intervention, assessment of outcomes and generation of hypotheses; see Higgins & Green, 2008):

  o RCT - randomized controlled trial

  o QRCT - quasi-randomized controlled trial (i.e., participants are allocated by means such as alternate allocation, person's birth date, the date of the week or month, or alphabetical order)

  o NRCT - non-randomized controlled trial (i.e. participants are allocated by other actions controlled by the researcher)

- Non-randomized studies (includes truly observational studies where the use of intervention has occurred in the course of usual decisions)

  o NRS - the allocation is not controlled by the researcher and there is a comparison of two or more groups of participants. Participants are allocated

by means such as time differences, location differences, decision-makers, or policy rules.

Non-randomized studies must demonstrate pre-treatment group equivalence via matching, statistical controls, or evidence of equivalence on key risk variables (e.g., parental education or socioeconomic status) and participant characteristics.

REFERENCES

Bascia, N., & Fredua-Kwarteng, E. (2008). *Class size reduction: What the literature suggests about what works.* Toronto: Canadian Education Association.

Biddle, B. J., & Berliner, D.C. (2002). Small class size and its effects. *Educational Leadership*, *59*(5), 12-23.

Finn, J. D. (2002). Small classes in American schools: Research, practice and politics. *Phi Delta Kappan*, *83*(7), 551-560.

Finn, J. D., & Achilles, C. M. (1990). Answers and questions about class size: A statewide experiment. *American Educational Research Journal*, *27*(3), 557-577.

Glass, G. & Smith, M. L. (1979). Meta-analysis of research on class size and achievement. *Educational Evaluation and Policy Analysis*, *1*, 2-16.

Grissmer, D. (1999). Conclusion: Class size effects: Assessing the evidence, its policy implications and future research agenda. *Educational Evaluation and Policy Analysis*, *21*(2), 231-248.

Hanushek, E. (1989). The impact of differential expenditures on school performance. *Educational Researcher*, *18*(4), 45-62.

Hanushek, E. (1999). Some findings from an independent investigation of the Tennessee STAR experiment and from other investigations of class size effects. *Educational Evaluation and Policy Analysis*, *21*(2), 143-165.

Hanushek, E. A. (2003). The failure of input-based schooling policies. *Economic Journal 113*(1), F64-F98.

Hedges, L. V., Laine, R. D., & Greenwald, R. (1994). Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educational Researcher, 23*(3), 5-14.

Hedges, L. V. & Olkin, I. (1985) *Statistical methods for meta-analysis.* New York: Academic Press.

Hedges, L. V. & Stock, W. (1983). The effects of class size: An examination of rival hypotheses. *American Educational Research Journal, 20*(1), 63-85*.*

Higgins, J. P. T., & Green, S. (Eds.) (2008). *Cochrane handbook for systematic reviews of interventions.* The Cochrane Collaboration.

Krueger, A. B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics*, *114*(2), 497-532.

Krueger, A. B. (2003). Economic considerations and class size. *The Economic Journal, 113*(February), F34–F63.

Nye, B., Hedges, L. V., & Konstantopoulos, S. (1999). The long-term effects of small classes: A five-year follow-up of the Tennessee class size experiment. *Educational Evaluation and Policy Analysis, 21*(2), 127-142.

AUTHOR(S) REVIEW TEAM

Lead reviewer:

Name:	Trine Filges
Title:	Senior Researcher
Affiliation:	SFI-Campbell
Address:	Herluf Trollesgade 11
City, State, Province or County:	Copenhagen
Postal Code:	1052
Country:	Denmark
Phone:	45 33480926
Email:	tif@sfi.dk

Co-author(s):

Name:	Christoffer Scavenius Sonne-Schmidt
Title:	Researcher
Affiliation:	SFI-Campbell
Address:	Herluf Trollesgade 11
City, State, Province or County:	Copenhagen
Postal Code:	1052
Country:	Denmark
Phone:	45 33480971
Email:	css@sfi.dk

Name:	Tine Nielsen
Title:	Senior Researcher
Affiliation:	SFI-Campbell
Address:	Herluf Trollesgade 11
City, State, Province or County:	Copenhagen
Postal Code:	1052
Country:	Denmark
Phone:	45 33480819
Email:	tnn@sfi.dk

Name:	Anne Marie Klint Jørgensen
Title:	Librarian/Information Specialist
Affiliation:	SFI-Campbell
Address:	Herluf Trollesgade 11
City, State, Province or County:	Copenhagen
Postal Code:	1052
Country:	Denmark
Phone:	45 33480868
Email:	amk@sfi.dk

ROLES AND RESPONSIBLIITIES

- Content: Christoffer Scavenius Sonne-Schmidt and Tine Nielsen

- Systematic review methods: Trine Filges

- Statistical analysis: Christoffer Scavenius Sonne-Schmidt, Tine Nielsen and Trine Filges

- Information retrieval: Anne Marie Klint Jørgensen

SUPPORT

No support is requested at this time.

FUNDING

Internal funding: SFI-Campbell

External funding: None

POTENTIAL CONFLICTS OF INTEREST

None known

PRELIMINARY TIMEFRAME

Approximate date for submission of Draft Protocol: 30 August 2012

DECLARATION

Authors' responsibilities

By completing this form, you accept responsibility for preparing, maintaining, and updating the review in accordance with Campbell Collaboration policy. The Coordinating Group will provide as much support as possible to assist with the preparation of the review.

A draft protocol must be submitted to the Coordinating Group within six months of title acceptance. If drafts are not submitted before the agreed deadlines, or if we are unable to contact you for an extended period, the Coordinating Group has the right to de-register the title or transfer the title to alternative authors. The Coordinating Group also has the right to de-register or transfer the title if it does not meet the standards of the Coordinating Group and/or the Campbell Collaboration.

You accept responsibility for maintaining the review in light of new evidence, comments and criticisms, and other developments, and updating the review at least once every three years, or, if requested, transferring responsibility for maintaining the review to others as agreed with the Coordinating Group.

Publication in the Campbell Library

The support of the Coordinating Group in preparing your review is conditional upon your agreement to publish the protocol, finished review and subsequent updates in the Campbell Library. Concurrent publication in other journals is encouraged. However, a Campbell systematic review should be published either before, or at the same time as, its publication in other journals. Authors should not publish Campbell reviews in journals before they are ready for publication in the Campbell Library. Authors should remember to include the following statement in reviews that are published elsewhere: "This is a version of a Campbell review, which is available in The Campbell Library".

I understand the commitment required to undertake a Campbell review, and agree to publish in the Campbell Library. Signed on behalf of the authors:

Form completed by: Trine Filges   **Date: 10 October 2012**