

GLOBALISING THE CLASS SIZE DEBATE: MYTHS AND REALITIES

Kevin Watson
Boris Handal^a
Marguerite Maher
Erin McGinty

University of Notre Dame, Australia

ABSTRACT: *Public opinion reflects a 'common sense' view that smaller classes improve student academic performance. This review reveals that the 'class size' effect of increased academic performance, although significant for disadvantaged students and students in the very early years of schooling, does not necessarily transfer to other student groups. Moreover, the literature indicates there are other more cost-effective variables that enhance student learning outcomes such as those associated with teacher quality. Internationally, large-scale interventions concluded that systematic class size reductions were more resource intensive requiring more personnel, training and infrastructure. From the large quantitative studies of the 1980s to the more qualitatively focused research in the last decade, there is a now an understanding that class size reductions function to provide opportunities for more student-focused pedagogies and that these pedagogies may be the real reason for improved student academic performance. Consequently, the impact of class size reductions on student performance can only be meaningfully assessed in conjunction with other factors, such as pedagogy.*

KEYWORDS: *class size, learning outcomes, educational policy, schooling, pedagogy*

Introduction

The vast majority of research about the effectiveness of class size reduction on increasing student academic achievement, at first glance, confirms what many members of the general public would regard as 'common-sense'. Namely, that class size reductions improve student performance (Finn and Achilles, 1999; Schanzenbach, 2007; Konstantopoulos, 2009). This is because research findings are evaluated simultaneously at two levels. The first is at the level of the research itself while the second is in terms of 'common-sense' expectations (Hanushek, 1998; Pritchard, 1999). This means that the media, teachers and members of the general public as well as researchers read and interpret research with a view to confirm their 'common sense' expectations. It is only when research findings clearly contrast with what is expected that they are questioned or alternative interpretations sought. Consequently, when research findings about class size are criticised (Hanushek, 1989; Hedges, Laine and Greenwald, 1994; Hanushek, 1996; Prais, 1996; Hoxby, 2000) the reasons for the criticisms may be 'subconsciously' devalued or questioned in line with 'common-sense' expectations.

Historical Background

Towards the end of the 1990s and into the 2000s researchers were beginning to strongly voice their opinion that class size reductions should not be expected to improve student academic performance and that the relationships between class size reductions and student improvement was relatively weak. They began to re-examine, re-think and therefore re-interpret the evaluations of previous studies.

In particular, there were three studies that have investigated the relationship between class size reductions and student academic performance that have survived multiple evaluations over many years and so make these studies very credible and also very influential in all aspects of the debate about class size reductions. Together, these three studies have profoundly influenced international

^a correspondence can be directed to boris.handal@nd.edu.au

thinking about the impact of class size on student achievement so a thorough understanding of what they do say and what they do not say is essential. The three studies are the:

- a) Student-Teacher Achievement Ratio (STAR) project;
- b) Student Achievement Guarantee in Education (SAGE) project; and
- c) California Class Size Reduction Program (CSR).

Although these three studies were conducted to find out if reductions in class size improved student achievement, they were not designed to show how or why smaller class sizes improved student achievement. Nevertheless, knowledge of these three studies is essential to contextualise a rigorous interrogation of the role class size plays on student achievement.

Student-Teacher Achievement Ratio

The best known of these studies is the Student-Teacher Achievement Ratio (STAR) project, a four year experimental study conducted from 1985 to 1989 (Finn and Achilles, 1990; Mosteller, 1995; Konstantopoulos, 2009). The project compared the academic achievement of small classes of 13 to 17 students with the 'regular' class size of 22 to 26 students across kindergarten, first, second and third grade in the American state of Tennessee (Finn and Achilles, 1990; Mosteller, 1995; Konstantopoulos, 2009). The project also compared the academic achievement of the 'regular' classes with and without an instructional aide to assist the classroom teacher. None of the teachers participating in the study received professional learning about teaching strategies specific to the class size they were teaching (Konstantopoulos, 2009). The reliability of the project findings was enhanced because the experimental design included a control (Finn and Achilles, 1990; Mosteller, 1995; Konstantopoulos, 2009).

The study was large scale. It included 79 schools, over 300 classrooms and about 7,000 students (Pritchard, 1999). Class sizes were maintained within their size limits for the four years of the study using an elaborate, predetermined process. The reliability of the findings was further enhanced by randomly assigning the teachers and students to each of the three different types of classrooms (Finn and Achilles, 1990; Mosteller, 1995; Pritchard, 1999; Konstantopoulos, 2009). All schools and classrooms made a similar contribution to the study by having, and maintaining, at least one class in each of the three classroom categories (Pritchard, 1999).

The findings showed that students in the smaller classes outperformed students in the larger classes whether the larger classes had an aide or not (Finn and Achilles, 1990; Pritchard, 1999; Konstantopoulos, 2009). Overall, the STAR project found that students in smaller classes outperformed students in larger classes on both standardised and curriculum-based tests. This was the case irrespective of whether the students were white or members of minority groups or whether they lived in the inner-city, urban, suburban or rural environments (Word et al., 1990). Specifically, the achievement effect of smaller classes on minority students was twice that for white students at the beginning of the experiment and remained that way over time (Pritchard, 1999; Konstantopoulos, 2009).

The Lasting Benefits Study was a follow-up study of the STAR Project. It began in 1989 to provide additional evidence about the effects of class size reduction on academic performance (Pritchard, 1999). Specifically, the Lasting Benefits Study was designed to find out if the benefits of small class size persisted when the students were returned to regular size classes (Pritchard, 1999). The findings showed that in the fourth grade, the original kindergarten students from the smaller classes outperformed students from the larger classes in all academic subjects (Achilles et al., 1993; Finn and Achilles, 1999; Pritchard, 1999; Konstantopoulos, 2009). The same students were better behaved (effort, initiative and disruptiveness) than students from the larger classes. The study also showed that through to the eighth grade, the original year four students from the smaller classes were still academically outperforming students from the larger classes. However, the level of outperformance decreased each year (Pritchard, 1999).

Project Challenge was the third phase of Project STAR and sought to implement its findings. It began in 1990 with 16 of Tennessee's poorest school districts (Pritchard, 1999). Project Challenge phased in small classes from kindergarten to third grade. The project was evaluated by ranking school districts based on their performance in a state-wide achievement test (Pritchard, 1999). The 16 school districts that participated in Project Challenge moved from the bottom of the performance table to the middle in reading and mathematics for second grade classes (Nye, Hedges & Konstantopoulos, 1999; Pritchard, 1999). Project Challenge also found that the number of students who repeated a grade was reduced in project classes (Pritchard, 1999).

Overall, the three Tennessee studies showed that small classes have an advantage over larger classes in school performance in the early primary grades (Finn and Achilles, 1990) and provided strong evidence regarding the positive effects of class size reductions. However, Hanushek (1998) summarised the whole project with a single insight. He states that at the end of kindergarten the children in the smaller classes scored higher than those in larger classes. He then agrees that the difference between the scores was maintained over the next three years. The fact that it did not widen, Hanushek (1998) argues, is clear evidence that the advantage of smaller class size ceased after kindergarten. Kindergarten was the only year in which class size reductions improved student academic performance.

Hanushek (1998) argues that significant reductions in class size, about one third less than 'normal' class size, may result in improved achievement for kindergarten students. He states that reductions to 18 or 20 students in kindergarten classes may not be as effective since this order of reduction is similar to the decline in the student to teacher ratio achieved between 1950 and the year data were collected for Project STAR. He suggests that no reasonable interpretation of the findings from Project STAR indicate that reductions in class size would improve student academic performance beyond kindergarten.

Student Achievement Guarantee in Education

The Student Achievement Guarantee in Education (SAGE) program began in 1996. The program was based in Wisconsin and phased in class size reductions from Kindergarten (1996-7) to grade three (1998-9) over three years in school districts where students came from low-income families (Molnar et al., 1999; Pritchard, 1999). The program lasted a total of five years and reduced class size to a student to teacher ratio of no more than 15:1 (Molnar et al., 1998; Pritchard, 1999). This does not mean that classes consisted of 15 students per teacher. Classrooms ranges from 'normal' arrangements of 15 students per teacher, through teams of two teaching 30 students to more elaborate arrangements reflecting site constraints and teacher availability (Molnar, Smith and Zahorik, 2000). When the evaluation of the program began in the 1997-8 school year there were 30 schools in 21 districts participating in the program with 14 schools across 7 districts taking part in the evaluation (Pritchard, 1999).

The SAGE program evaluated student achievement at the beginning and end of first-grade and second-grade. The results were compared with those of similar schools with 'normal' sized classes. The comparison schools were similar in terms of grade levels, enrolment, socioeconomic status, ethnicity, race and gender (Pritchard, 1999). It was found that SAGE first-grade students performed consistently better than students in schools with 'normal' size classes.

The achievement difference between SAGE and comparison schools was greatest at the end of the first test period (first-grade) and then, although still greater, the difference decreased in subsequent years (Molnar et al., 1999; Pritchard, 1999). The achievement gap between white and African-American first-grade students was less for SAGE schools compared with students in schools with 'normal' size first-grade classes (Pritchard, 1999). Although the achievement for second-grade SAGE students remained higher than for students in schools with 'normal' size classes (Molnar et al., 2000; Graue and Oen, 2009), the difference did not increase over time but remained the same (Blatchford, Goldstein and Mortimore, 1998; Grissmer, 1999; Pritchard, 1999).

It is important to acknowledge that as well as a reduction in class size, the SAGE schools simultaneously implemented a rigorous academic curriculum as well as professional learning and after school programs (Molnar et al., 1998; Krueger, 1999). The researchers responsible for the program indicated that these additional components of the program did not have time to impact student achievement. However, an influence cannot be ruled out (Blatchford, Goldstein and Mortimore, 1998; Grissmer, 1999). The study was also criticised because SAGE schools were not similar enough to comparison schools and the variation in what constituted a student to teacher ratio of 15 to 1 was too varied and often inconsistent across the SAGE and comparison schools (Blatchford, Goldstein and Mortimore, 1998; Grissmer, 1999; Pritchard, 1999). Other criticisms were that both teachers and students were not randomly allocated to classes and that the study did not match pair SAGE and comparison schools (Pritchard, 1999). Researchers also criticised the SAGE program because classroom composition changed in both SAGE and comparison schools annually (Blatchford, Goldstein and Mortimore, 1998; Pritchard, 1999). Also, the SAGE program did not continue the evaluation beyond fourth grade. However, the greatest criticism of the program was that increased achievement of SAGE students was due not only to reduced class sizes but to the increased attention paid to students as a result of the rigorous academic curriculum, professional learning and after school programs (Molnar et al., 1998; Krueger, 1999).

California Class Size Reduction Program

The California Class Size Reduction Program (CSR) attempted to implement the findings of the STAR project across the state of California (Mitchell and Mitchell, 1999; Stecher and Borhnstedt, 2000). The program was phased in over four years. However, there was no longitudinal dimension to the program.

The CSR was evaluated by Mitchell and Mitchell (1999) two years after its implementation and involved over 80 schools across eight school districts in Southern California. Reading, language and mathematics subtests were analysed from the Stanford Achievement Test (SAT) as well as more than 30 variables related to student demographics, classrooms, teachers and schools. The demographic variables included gender, student poverty, ethnicity and the language spoken at home (Mitchell and Mitchell, 1999). When these variables were controlled for, it was found that students in reduced class sizes made small but statistically significant gains in achievement when compared with students in 'normal' size classes. Reduced class size was considered to be not more than 20 while a class size over 20 was considered to be 'normal' (Mitchell and Mitchell, 1999).

The evaluation concluded that other factors may have contributed to gains in achievement (Mitchell and Mitchell, 1999). The other factors included new requirements for teacher education programs, revisions to bilingual programs, new curriculum frameworks and materials, new statewide tests and a new performance accountability system (Mitchell and Mitchell, 1999). None of these factors were taken into account during the implementation of the program and any or all of them could have influenced the results of the program and the findings of the evaluation.

Stecher and Borhnstedt (2000) also evaluated the study and found small increases in academic achievement by students in small classes. However, closing the gap between majority and minority groups was found to be not as significant. In addition, Stecher and Borhnstedt (2000) criticised the design of the program because it did not include a control group as part of its quasi-experimental design. They also found no basis for how 20 students was considered to be the difference between a small class and a 'normal' class.

Researchers (Mitchell and Mitchell, 1999; Stecher and Borhnstedt, 2000; Funkhouser, 2009; Jepsen and Rivkin, 2009) found that increased student attention in class and the attention paid by teachers to students as a result of the smaller class sizes contributed to increased student achievement. Specifically, teachers in the smaller classes spent more time teaching, were able to address individual student differences, had more interactions with each individual student and were able to minimise class disruptions compared with classes of 'normal' size. The question remains,

could these reasons for increased student achievement have been achieved in ways other than reducing class size?

In addition to the above, an evaluation by Jepsen and Rivkin (2009) found that the statewide scale of the Californian project led to a shortage of qualified teachers. Paradoxically, the teacher shortage was most acute in low income school districts. To overcome this problem, teachers were employed who were less qualified, lacked teaching experience and who were in other ways less suitable to teach at-risk, low income students. Consequently, Jepsen and Rivkin (2009) qualified their evaluation by indicating that their findings did not reflect a study of only one variable, namely, a reduction in class size. The evaluation conducted by Funkhouser (2009) concluded that the increases in student achievement may have actually been underestimated because of the reduced teacher qualifications. However, Funkhouser (2009) also pointed out that any increase in student achievement was also likely to be attributed to the other programs that were simultaneously implemented across the state.

Overall, the studies carried out by Mitchell and Mitchell (1999) and by Stecher and Borhstedt (2000) found that the large scale of the program meant it was extremely expensive to implement. The expenses were a result of the increase in the number of teachers employed and the additional classrooms required. The researchers found it difficult to attribute the small gains in student achievement to reduced class size because of the problems associated with the quasi-experimental design of the program, the lack of a rationale for deciding that the difference between reduced and 'normal' class size was twenty students and because there were no base-line data for student achievement established using pre-tests before the program was implemented. However, the evaluations carried out by both Mitchell and Mitchell (1999) and Stecher and Borhstedt (2000) cited the lack of the random assignment of students to the treatment groups as the greatest threat to internal validity. Therefore, the various teams of researchers concluded that the Californian CSRFP did not replicate the STAR project as a controlled randomised longitudinal study.

Recent Research About Reductions in Class Size

During the 1990s research on class size reduction was dominated by the STAR (Tennessee), SAGE (Wisconsin) and CSRFP (California) studies with an emphasis on quantitative analyses. The following decade was influenced by research studies that included both quantitative and qualitative methodologies seeking to articulate more specifically the relationship between class size and attainment while focusing on other contextual mediating variables. Class size reduction seems to be a variable that, while it can be investigated in its own right, can only be understood in the context of other school reforms that influence academic achievement such as new views about curriculum, emerging pedagogies, the introduction of standards-based testing and teaching quality. Like any other educational analysis, any class size reduction investigation must also take into account a broad range of socio-cultural characteristics exogenous to the school such as socio-economic disadvantage and ethnicity.

An emerging theme from a review of the literature after the year 2000 is that class size reductions, in general, have an effect on the very early years of schooling as well as for disadvantaged groups. Studies that draw this conclusion make more use of qualitative methods in the quest for contextualising and understanding class size reductions within a broader framework including not only a uni-dimensional relationship to standardised achievement but also in terms of group dynamics, school subjects, interactions in class and / or socio-cultural variables.

This next section presents class size research reported mainly after the year 2000 from English speaking countries such as Australia, the United Kingdom, Canada, the United States and New Zealand.

Australia

Unfortunately, there is no published research about class size reductions based on Australia data. The only reference to Australian class size data comes from the Third International Mathematics and Science Study (TIMSS) which reported that the actual class size (or grade-average class size) for

13-years-old mathematics classes is 26.962 students (Woßmann and West, 2006). The other TIMSS sample countries, with average mathematics class sizes provided in brackets, are: Belgium-French (20.087), Belgium-Flemish (20.330), Canada (27.813), Czech Republic (25.367), France (25.567), Greece (28.555), Hong Kong (40.611), Iceland (20.136), Japan (36.334), Korea (50.513), Portugal (25.645), Romania (27.436), Scotland (26.190), Singapore (32.493), Slovenia (24.215), Spain (28.551) and the United States (25.909).

United Kingdom

Relevant among those studies using mixed mode research identifying class size effects is the work conducted by Peter Blatchford and his associates at the Institute of Education, University of London. Blatchford et al. (2001) investigated the links between class size and within class groupings with 6,672 Reception, Year 2 and Year 5 classes in 331 United Kingdom schools. The researchers suggested that class size had an effect on the size and number of groups within a class which in turn influenced student learning experiences. In general, the authors suggested that class size was inversely proportional to group size but directly proportional to number of groups. For example, larger classes in the study tended to work in groups while smaller classes seemed to adopt a whole class approach. Furthermore, it was found that children in larger classes worked and were taught in group sizes above the average.

A year later the study was extended to Year 4 to Year 6 students aged 7-11 years from over 200 schools (Blatchford et al., 2007). The researchers found more individualised teaching in smaller classes but, in contrast to the previous study, teachers used more whole class teaching in both larger and smaller classes. The authors suggested that this might be due to the pressure of covering the required content particularly in mathematics, science and literacy as the students and curriculum moved from primary towards secondary education. The authors also argued that in smaller classes opportunities were not always available for group work because of the less differentiated nature of the curriculum. The researchers argued that teachers in smaller classes still used pedagogies appropriate for larger classes and this was corroborated by previous research (Betts and Shkolnic, 1999). They said, "there is no guarantee that smaller classes will automatically lead to more productive work in groups" (p. 169). The study also revealed that there were more teacher-student interactions in smaller classes. Larger classes seemed to pose difficulties with classroom management, student passivity, group size and teaching / learning space in the classroom. In addition, teachers expressed their concerns about the larger amount of marking in large classes and a strong belief that small classes facilitate reaching individual students.

Blatchford et al. (2002) investigated class size effects in the subject English with students in Key Stage 1, made up of Reception, Year 1 and Year 2 students (4-7 year olds). The study focused on the achievement of 9,330 students in 220 schools. The authors concluded that there was a positive relationship between class size and initial achievement particularly for low achievers and for students entitled to free school meals (low family income levels). However, there was no relationship found between class size and student achievement in mathematics. In this study smaller classes ranged from 18 to 25 students.

Through systematic classroom observations the effect of class size on student classroom engagement and teacher-student interaction was investigated in 49 randomly selected primary and secondary schools (Blatchford, Bassett and Brown, 2008). The findings supported the belief that in small classes low performing students tended to receive more teacher attention in the form of active interaction with the teacher for all levels of achievement. Similarly, smaller classes provided an environment where the teacher-student interaction increased compared with larger classes. In the study there were positive interactions between attainment groups and class size for secondary but not for primary students.

Class size also had an effect on student off-task behaviour particularly for lower performing groups in both primary and secondary contexts with off-task behaviour increasing as classes

become larger. The variable 'student focus on teacher' also increased as class size decreased for all attainment levels. It was also noted that the amount of teaching increased with class size. For example, in secondary schools "An increase in class size of 5 pupils was associated with the odds of teacher teach [teacher instruction] increasing by 8%" (Blatchford, Bassett and Brown, 2008, p. 17). Blatchford, Bassett and Brown (2008) also found that classroom management incidences such as disruptive classroom behaviour increased with class size for low and medium performing primary students and for medium performing secondary students. In their analysis of this and other studies Konstantopoulos and Chung (2009) concluded that, in general, class size effects go beyond the early primary years and that this is consistent with previous research.

Overall, the research from the United Kingdom on the relationship between class size and student achievement found that reductions in class size increased student performance although this was not always the case. It can be concluded that increases in student performance were related to implementing pedagogies that provided more opportunities for individual and small group work. In addition, greater academic performances were achieved for lower performing students and for students from lower socio-economic backgrounds particularly in the earlier years of schooling.

Canada

Two studies in Alberta and Ontario, Canada revealed improvements in student achievement as a result of class size reduction programs. The Alberta study (Bascia and Fredua-Kwarteng, 2008) investigated the effect of class size on achievement in first grade 'high-needs' schools in Edmonton. In partnership with the public education sector, classes of 15 or fewer students were created. A distinctive feature of the Alberta study was the provision of professional learning programs that were implemented with the class size reduction program. The project reported substantial improvement in academic scores after five months. A reduction in disruptive behaviour was also reported. Achievement was measured using the Canadian Test of Basic Skills (CTBS), Developmental Reading Assessment (DRA) and the Highest Level of Achievement test (HLA). Teachers completed questionnaires and were interviewed to assess the effects of integrating reading, writing and speaking. The researchers suggested that class size reduction programs should target specific student populations, be accompanied by professional learning programs and support, be implemented gradually starting from the early years of schooling and have sustainable funding.

The Ontario class size reduction program aimed to ensure that 90% of all primary school classes had no more than 20 students and that 100% of those classes had 23 or fewer students. Funding was provided to sustain new teachers as well as infrastructure as early as 2004. The participation goal was met in the 2009 / 2010 school period. An initial assessment of the program in 2007-2008 (Bascia et al., 2010) using qualitative methods such as interviews and observations revealed positive results in teaching and learning. This included more meaningful communication in class, more frequent one-to-one work with students, increased group work as well as a greater diversification in teaching strategies particularly in literacy. It is noteworthy that the class size reduction program was accompanied by professional development initiatives particularly in literacy. Teachers reported their satisfaction in being able to monitor students more closely, providing more personalised delivery of the curriculum and being able to run open-ended projects with their students. They also appreciated that smaller classes facilitated dealing better with disadvantaged students which was a focus of the program. Less disruptive behaviour was also reported as another benefit of the class reduction project. As one teacher reported, "I can stay on top of everything more. I know how they (students) are doing in all areas. There is no place to 'hide' in a small class" (ibid, p. 74). More classroom space became available as a result of fewer students in a class, which helped provide a more suitable environment to conduct group activities. However, the report noted that some teachers could have taken greater advantage of smaller classes to employ a more differentiated approach to teaching and learning and to help students from disadvantaged backgrounds.

The Canadian studies affirmed the finding that smaller class sizes provide opportunities to implement different pedagogies and that it is the pedagogy employed that facilitates increased student / teacher communication, one-to-one work, group work and an increased diversity of teaching strategies that is, in turn, responsible for increased student performance. The shift in emphasis from quantitative to the inclusion of qualitative studies provides the opportunity for researchers to investigate the reasons for and an understanding of increased student performance rather than simply noting that improved achievement occurs sometimes yet not at others. The Canadian studies also confirmed that improved student achievement is increased in 'high needs' schools, when the intervention occurs in the earlier years of schooling and when accompanied by professional learning. Consequently, it is becoming increasingly clear that when class size reduction programs are implemented, their success in improving student achievement seems to be related to the opportunity they provide to implement different, more learner focused pedagogies. The Canadian studies also indicate that it may be the professional learning programs that were part of the overall initiative that helped teachers to become aware of and then implement pedagogies that promote learning in smaller classes.

United States

In 1999, as a result of the momentum gained from earlier class size reduction initiatives, particularly as a result of the success of the Tennessee STAR project, Congress approved a federal program to facilitate class size reductions to no more than 18 students in grades 1 to 3. An additional 110,000 new teacher graduates were employed (Scudder, 2002) and by 2005, 24 states had mandated or implemented class size reduction programs (Education Commission of the States, 2005).

Most of the United States studies on class size reduction revealed increases in academic performance although some contradictory findings were reported. Milesi and Gramoran (2006) analysed standardised test data from the Early Childhood Longitudinal Study – Kindergarten Class of 1998-99 and found no statistically significant relationship between class size and student achievement in either mathematics or reading. This was a national study which included 21,260 children enrolled in approximately 1,000 kindergartens and grouped according to three different class sizes: small (17 or fewer students), regular (between 18 and 23 students) and large (24 or more students). The study found no class size effects based on race or ethnicity and economic and academic background.

Funkhouser (2009) followed the academic achievement of Californian second grade students from 135 schools that had participated in class size reduction programs from kindergarten. A comparison group was formed using a kindergarten cohort that had, at no stage, been part of a class size reduction program. Funkhouser (2009) found a class size reduction effect in reading and mathematics but not in language and spelling. Similarly, the impact of class size reduction programs for Latino and non-English speaking background students was found to be minimal with the greatest variation in test scores attributable to student demographic and economic background and to teacher characteristics. Funkhouser (2009) concluded that compared to demographic, economic and teacher variables, the influence of class size reductions on student achievement although positive was relatively immaterial.

In a similar study Stecher, McCaffrey and Bugliari (2003) examined the effect of class size reduction programs with third grade students from 1,918 Californian schools. The sample consisted of students who had entered kindergarten in 1995-96, 1996-97 and 1997-98. As a consequence of the timing of the federal class size reduction program, the students across this range of entry into kindergarten experienced up to four different class size reduction programs from kindergarten to grade three. For example, third grade students who entered kindergarten in 1995-96 had no exposure to class size reduction interventions while those who enrolled in kindergarten in 1997-1998 had participated in up to four class size reduction programs. In general, Stecher, McCaffrey and Bugliari (2003) found no relationship between student academic achievement and class size sequence of exposure.

The findings from the Stecher, McCaffrey and Bugliari (2003) study contrast with previous research (Stecher, McCaffrey and Burroughs, 1999; Stecher et al., 2000). Consequently, Stecher, McCaffrey and Bugliari (2003) argue that judging class size reduction interventions by single year increments is inconclusive. They suggest that class size reduction programs should be evaluated in terms of their cumulative effects.

Maasoumi, Millimet and Rangasaprad (2005) found no overall class size reduction effects across a national representative sample of tenth and twelfth grade student academic scores in reading, social studies, mathematics and science. However, when class size was examined more specifically they found that reducing class size for classes with more than 20 students did not greatly influence attainment scores. However, when class size was reduced below 20, greater effects were found. In classes with less than 20 students test scores increased for students who were below the median test score but decreased for students above the median test score.

Interestingly, two studies in different states using the same statistical methods showed conflicting results. Working with fourth and sixth grade student scores drawn from data across the state of Connecticut, Hoxby (2000) found no class size reduction impact on achievement while Cho, Glewwe and Whitler (2010) working with grade three and four Minnesota students found a small class size reduction effect. The picture that is building about recent research concerning class size reduction programs is that findings are inconsistent and seem to be cohort dependent rather than universally transferable.

Research from North Carolina revealed class size reduction effects on achievement. In 2001 the North Carolina government approved the funding of a class size reduction program targeting kindergarten to grade three classes in 36 high priority schools. The project planned to reduce class sizes to a maximum of 15 students (Casbon, DeMeester and Nalley, 2002). In 1991 Burke County schools in North Carolina had already begun a class size reduction program. By 2001 all grade one to three classes in its 17 elementary schools, mostly rural, had reduced their class sizes to below 20. Classroom observations throughout this period showed that teachers in the smaller classes spent more time communicating with students compared with those teaching regular size classes. End-of-grade test score data in mathematics and reading across grades three to seven were collected for students in small classes and compared with those from regular sized classes. Casbon, DeMeester, and Nalley (2002) found an increase in the performance of students in small classes and this level was maintained in subsequent years lasting up to grade nine for English.

In North Carolina, Scudder (2002) investigated the effect of the federal class size reduction program in 23 schools in Wake County during the 1999-2000 school year. Schools were selected based on the number of disadvantaged students. The sample included students from grades one to three who receive free or reduced-price lunches as well as students who were achieving below their grade level. The program hired 23 additional teachers (averaging one teacher per school) to create new classes so that all classes were in the range of from 23 to 18 students. A total of 2,474 students benefited from the class size reduction program. Pre and post literacy county-based end-of-grade tests were used at the beginning and end of the program. To facilitate data analysis, a comparison group was randomly formed from students not affected by the intervention but who were from the same pool of schools. The comparison group reflected the same demographics as the intervention group. Scudder (2002) found that students participating in the intervention increased their academic performance more than those in the comparison group in first and second grade. The increase in performance was quantitatively similar to that of the STAR project. The findings also showed that students who received free or reduced price lunches demonstrated significantly higher improvement compared with those in the comparison group.

The recent United States experience with class size reduction programs supports the emerging view that such programs are seen to increase student academic performance when reductions occur below a class size of 20 and when disadvantaged or low performing students are targeted. More importantly, these more recent studies show that compared with demographic, economic and teacher effects, the influence of class size reductions is not as great. These studies also show that

the reasons for the effects are often associated with pedagogy. However, overwhelmingly, recent United States studies show that research about the influence of class size reduction programs on improving student academic outcomes is inconclusive with many studies showing little or no effect.

New Zealand

Hattie (2009) in his synthesis of over 800 meta-analyses related to student achievement, argues that it is not difficult to find evidence on both sides of the discussion about whether a reduction in class size leads to enhanced classroom learning or not. He discusses the nuanced arguments put forward by researchers in terms of what constitutes a small class, the interpretation of findings and the interaction of class size reductions with other factors such as teacher quality. Hattie (2009) attempts to identify the reasons for the positive effect of class size reductions and in so doing positions its influence in terms of other, interrelated influences. He argues that reduced class size facilitates individualised instruction; higher quality instruction; greater scope for innovation and student-centred teaching; increased teacher morale; fewer disruptions; less student misbehaviour; and greater ease in engaging students in academic activities.

The bottom-line of Hattie's (2009) findings is that reducing class size has little influence on increasing student academic achievement because the pedagogical changes more strongly associated with improved student achievement in small classes are not implemented. In other words, teachers do not change their teaching practices because they are teaching smaller classes. They continue to teach as if they were teaching larger classes. It is not the size of the class that increases student academic performance. Rather, it is the teaching practice or pedagogy that a reduction in class size facilitates that improves student performance.

Hattie positions class size in a hierarchy where class size is one of a number of mediators that enable and then facilitate the potential for pedagogical change – pedagogy that facilitates a more student focused approach to teaching and learning. Hattie makes the point that, in isolation, class size is a small factor in increasing student academic performance. Class size reduction achieves its 'effect size' as an enabler of increased student achievement only when the desired pedagogies are implemented effectively. This takes both time and expertise which goes some way to explaining the inconsistencies associated with class size reduction interventions.

Snook et al. (2009) has made a number of criticisms about class size research. They argue that 'in-school' influences have been the main focus of class size research while the influence of social factors has largely been ignored. Social factors, including culture, influence student performance irrespective of class size and therefore the transfer of research findings across cultural divides may be problematic. The findings from United States studies, although they may be indicative of what could be the case in New Zealand and Australia are contextualised by different cultures and therefore different findings may result. Hattie's synthesis and subsequent interpretations were made from the perspective of being informed from culturally embedded New Zealand worldviews.

Conclusion

Class size reduction has a long history of interventions and associated research. At times this research has simultaneously been read and interpreted in terms of 'common sense' expectations as well as the research in its own right. Consequently, there is a general belief that small classes increase student academic performance almost irrespective of research findings while the reasons for such findings go unexplored. This is a strong statement to make. However, it is supported by both a careful and accurate interpretation of research results (Hanushek, 1998) and an interpretation of the findings from a number of studies over an extended period of time (Pritchard, 1999).

Essentially, more recent research sought to articulate the specific relationship between class size and student attainment while recognising that the relationship was complex and, as a consequence accommodated other mediating variables. The class size reduction variable seems to be a variable

that, while it can be investigated in its own right, can only be understood in the context of other school reforms that influence academic achievement such as new views about curriculum, emerging pedagogies, the introduction of standards-based testing and teaching quality.

One explanation for increases in student attainment not being explored beyond their relationship with class size reduction was the dominance of quantitative methods used to evaluate the early class size reduction interventions. This includes the three seminal United States studies (STAR, SAGA and CSR). However, a shift towards mixed research methods, which included qualitative data, revealed a variety of influences on student achievement related to class size reductions. In addition, different research designs and methodologies, sample characteristics and educational and cultural backgrounds may also explain the inconsistencies in research findings.

Mixed methods research also revealed the complex interaction of the class size reduction variable with other teaching and learning variables such as group work, classroom management, teacher-student interactions, teacher satisfaction and workload and student engagement. Interestingly, recently reported quantitative research in the United States shows that class size has an effect on student engagement in the middle school. These findings suggest that the cause-effect relationship between class size and achievement is mediated by other factors (Dee and West, 2011). More recent literature confirms these caveats (Galton and Pell, 2012; Cho, Glewwe and Whitley, 2012).

In the lack of more focused studies, further research is needed to articulate the full extent and complexity of this relationship and how far it extends to other mediating variables. Indeed, there is a whole spectrum of emerging research findings that identify these mediating effects as being integral to an understanding of the complex relationship between class size and student achievement.

Despite the vast amount of research about class size, there is no published research based on Australian data. The Third International Mathematics and Science Study (TIMSS) reported that across the world, class size effects were probably a result of teacher quality. Consequently, in an Australian context there are no real research findings that relate class size reductions to student academic achievement. There are only inferences derived from other countries and therefore other social, cultural and economic contexts. One of the emerging findings about class size research is that it tends to be socio-culturally and socio-economically dependent. Therefore, findings from other social, cultural and economic contexts cannot be assumed to apply to Australia, even though they may serve as reference points for further research.

This review has highlighted the fact that changes, over time, in the type of research methodologies employed have helped understand the relationships between class size reductions and student achievement by recognising the presence and importance of mediating variables. These mediating variables add a dimension of complexity to the relationship that goes some way in helping to understand the varied and inconsistent research findings drawn on and used in the debate about class size. Nevertheless, almost irrespective of the presence and influence of mediating variables, it is recognised that class size reduction is linked to increased academic performance for disadvantaged students (Krueger and Whitmore, 2002). The research also highlights the fact that class size reduction effects are greater in the earlier years of schooling, particularly kindergarten, and that this effect can flow into the early years of secondary education. What is not clear is whether maintaining smaller classes over this extended period of time helps to maintain the advantage gained in the early years or whether the advantage would persist without continuing with smaller classes. What is clear is that class reduction initiatives that have a strong staff professional learning component as part of their implementation appear to increase the potential for enhanced student achievement.

Perhaps the most significant outcome of this review is that it positions the relationship between class size reduction and student academic performance as dependent on pedagogy. Specifically, smaller classes do not achieve their expected performance outcomes if they are not accompanied by pedagogical changes that facilitate a more student-centred focus on teaching and learning. Teachers must be provided with learning opportunities that enable them to understand and use pedagogies that promote increased academic performance in small classes (Hattie, 2009). This might include a more differentiated learning environment that accommodates a diversity of student interests and

ability, the use of open-ended investigation techniques, an emphasis on collaborative group work, remedial assistance, and increased interactivity that discourages student passivity and includes the development of student help-seeking behaviours (Blatchford et al., 2007).

When the research on class size is viewed as a whole it is becoming increasingly clear that the 'common sense' view that class size reductions improve student academic performance need to be seen from the perspective that small class sizes may be an enabler and that the actual reasons for the perceived influence of small classes is they provide an opportunity for a more student-centred approach to teaching. It is this change in pedagogy that is really the 'common sense' component. This then raises the provocative question, can pedagogical change facilitate improved student achievement without the huge expense of employing more teachers, providing more classrooms and funding the increased infrastructure to support increased system management? Is it not better to spend far less money to achieve improved student performance by providing professional learning opportunities for teachers to learn how to teach in a more student-focused way? It is only then, when more student-focused pedagogies are in place that the actual effect of reducing class numbers can be evaluated. It may then be that class size reductions, in their own right, may enhance student performance.

References

- Achilles, C. M., Nye, B. A., Zaharias, J. B., & Fulton, B. D. (1993). *The Lasting Benefits Study (LBS) in grades 4 and 5 (1990-1991): A legacy from Tennessee's four-year (K-3) class-size study (1985-1989)*, Project Star, North Carolina Association for Research in Education, Greensboro, North Carolina.
- Bascia, N., Connelly, C., Flessa, J., & Mascal, B. (2010). *Ontario's Primary Class Size Reduction Initiative: Report on Early Implementation*. Canadian Education Association.
- Bascia, N., & Fredua-Kwarteng, E. (2008). *Class size reduction: What the literature suggests about what works*. Toronto: Canadian Education Association, 1-32. Available at <http://www.cea-ace.ca/sites/cea-ace.ca/files/cea-2008-class-size-literature.pdf>. [Accessed on 2 July 2013].
- Betts, J. R., & Shkolnic, J. L. (1999). The Behavioural Effects of Variations in Class Size: The Case of Math Teachers, *Educational Evaluation and Policy Analysis*, 21(2), 193-214.
- Blatchford, P., Goldstein, H., & Mortimore, P. (1998). Research on class size effects: A critique of methods and a way forward, *International Journal of Educational Research*, 29, 691-710.
- Blatchford, P., Baines, E., Kutnick, P., & Martin, C. (2001). Classroom contexts: Connections between class size and within class grouping. *British Journal of Educational Psychology*, 71, 283-302.
- Blatchford, P., Goldstein, H., Martin, C., & Browne, W. (2002). A study of class size effects in English school reception year classes. *British Educational Research Journal*, 28(2), 167-185.
- Blatchford, P., Russell, A., Bassett, P., Brown, P., & Martin, C. (2007). *The effect of class size on the teaching of pupils aged 7-11 years*. *School Effectiveness and School Improvement*, 18(2), 147-172.
- Blatchford, P., Bassett, P., & Brown, P. (2008). *Do low attaining and younger students benefit most from small classes? Results from a systematic observation study of class size effects on pupil classroom engagement and teacher pupil interaction*. New York: American Educational Research Association Annual Meeting, 1-29.
- Casbon, C., DeMeester, K. & Nalley, D. (Ed.) (2002). *Class Size Reduction. The Regional Educational Laboratory at SERVE*. SERVE: North Carolina.
- Cho, H., Glewee, P., & Whitley, M. (2010). *Do reductions in class size raise students' test score? Evidence from population variation in Minnesota's elementary schools*. Available at <http://faculty.apec.umn.edu/pglewee/documents/MNclasz3.pdf>. [Accessed on 2 July 2013].
- Cho, H., Glewee, P., Whitley, M. (2012). Do reductions in class size raise students' test scores? Evidence from population variation in Minnesota's elementary schools. *Economics of Education Review*, (31) 3, 77-95.

- Dee, T. S., & West, M. R. (2011). The non-cognitive returns to class size. *Educational Evaluation and Policy Analysis*, 33(1), 23-46.
- Education Commission of the States. (2005). *State Class-Size Reduction Measures*. Denver, Colorado: Education Commission of the States.
- Finn, J. D., & Achilles, C. M. (1990). Answers and questions about class size: A statewide experiment. *American Educational Research Journal*, 27(3), 557-577.
- Finn, J. D., & Achilles, C. M. (1999). Tennessee's class size study: Findings, implications, misconceptions. *Educational Evaluation and Policy Analysis*, 21(2), 97-109.
- Funkhouser, E. (2009). The effect of kindergarten classroom size reduction on second grade achievement: Evidence from California. *Economics of Education Review*, 28(3), 403-414.
- Galton, M. & Pell, T. (2012). Longitudinal effects of class size reductions on attainment: Results from Hong Kong primary classrooms. *International Journal of Educational Research*, 53, 360-369.
- Graue, M. E., & Oen, D. (2009). You just feed them with a long-handled spoon: Families evaluate their experiences in a class size reduction reform. *Educational Policy*, 23(5), 685-713.
- Grissmer, D.W. (1999). Class Size Effects: Assessing the Evidence, Its Policy Implications, and Future Research Agenda. *Educational Evaluation and Policy Analysis*, 21(2), 231-248.
- Hanushek, E. A. (1989). Expenditures, Efficiency, and Equity in Education: The Federal Government's Role. *American Economic Review*, 79(2), 46-51.
- Hanushek, E. A. (1996). A more complete picture of school resource policies. *Review of Educational Research*, 66(3), 397-409.
- Hanushek, E.A. (1998). The evidence on class size. *Public Testimony*, Washington, DC.
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Hedges, L. V., Laine, R., & 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.
- Hoxby, C. (2000). The effects of class size on student achievement: new evidence from population variation. *Quarterly Journal of Economics*, 115(4): 1239-1285.
- Jepsen, C., & Rivkin, S. (2009). Class size reduction and student achievement: The potential tradeoff between teacher quality and class size. *Journal of Human Resources*, 44(1), 223-250.
- Konstantopoulos, S. (2009). Effects of teachers on minority and disadvantaged students' achievement in the early grades. *Elementary School Journal*, 110(1), 92-113.
- Konstantopoulos, S. & Chung, V. (2009). What are long-term effects of smaller classes on the achievement gap? Evidence from the Lasting Benefits Study. *American Journal of Education*, 116(1), 125-154.
- Krueger, A. B. (1999). Experimental Estimates of Education Production Functions. *The Quarterly Journal of Economics*, 114(2), 497-532.
- Krueger, A. B., & Whitmore, D. M. (2002). Would smaller classes help close the black white achievement gap? In J.E. Chubb & T. Loveless (Eds.), *Bridging the achievement gap*. Washington DC: Brookings Institution Press. Lankford, pp. 11-46
- Maasoumi, E., Millimet, D. L., & Rangappasad, V. (2005). *Class size and educational policy: Who benefits from smaller classes?* Southern Methodist University. Available at <http://www.caen.ufc.br/~evea2007/maasoumi/sdclass.pdf>. [Accessed on 2 July 2013].
- Milesi, C. & Gramoran, A. (2006). Effects of class size and instruction on kindergarten achievement. *Educational Evaluation and Policy Analysis*, 28(4), 287-313.
- Mitchell, D. E., & Mitchell, R. E. (1999). *The impact of California's Class Size Reduction initiative on student achievement: Detailed findings from eight school districts*, California Educational Research Cooperative, University of California, Riverside: CA.
- Molnar, A., Smith, P., & Zahorik, J. (2000). *1999-2000 Evaluation Results of the Student Achievement Guarantee in Education (SAGE) Program*. Milwaukee, WI: Centre for Urban Initiatives and Research, University of Wisconsin

- Molnar, A., Smith, P., Zahorik, J., Palmer, A., Halbach, A., & Ehrle, K. (1999). Evaluating the SAGE program: A pilot program in targeted pupil-teachers reduction in Wisconsin, *Educational Evaluation and Policy Analysis*, 21(2), 167-177.
- Molnar, A., Percy, S., Smith, P., & Zahorik, J. (1998). *1997-98 Results of the Student Achievement Guarantee in Education (SAGE) Program*. Milwaukee, WI: University of Wisconsin-Milwaukee.
- Mosteller, F. (1995) The Tennessee Study of Class Size in the Early School Grades, *The Future of Children*, 5(2), 113-127.
- Nye, B. A., 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.
- Prais, S. J. (1996). Class size and learning: The Tennessee experiment – what follows? *Oxford Review of Education*, 22(4), 399-414.
- Pritchard, I. (1999). *Reducing Class Size, What Do We Know?* Available at http://www.ed.gov/pubs/ReducingClass/Class_size.html Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education. [Accessed on 2 July 2013].
- Schanzenbach, D. W. (2007). *What have we learned from Project STAR? Brookings Papers on Education Policy*, 205-208. Available at http://harrisschool.uchicago.edu/about/publications/working-papers/pdf/wp_06_06.pdf. [Accessed on 2 July 2013]
- Scudder, D. F. (2002). *An evaluation of the Federal class size reduction program in Wake County, North Carolina –1999-2000*. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA, April 1-5.
- Snook, I., Clark, J., Harker, R., O'Neill A., & O'Neill, J. (2009). Invisible Learnings? A commentary on John Hattie's "Visible Learning: A Synthesis of over 800 Meta-Analyses Relating to Achievement". *New Zealand Journal of Educational Studies*, 44(1), 93-106.
- Stecher, B. M., McCaffrey, D. M., & Burroughs, D. (1999). *Achievement*. In G. W. Bohrnstedt & B. M. Stecher (Eds.). *Class size reduction in California: Early evaluation findings, 1996–1998*. Palo Alto, CA: CSR Research Consortium.
- Stecher, B. M., & Bohrnstedt, G. W. (2000). *Class size reductions in California: The 1998-99 evaluation findings*. Sacramento, California: California Department of Education.
- Stecher, B. M., McCaffrey, D. F., Burroughs, D. Wiley, E., & Bohrnstedt, G. W. (2000). *Achievement*. In B. M. Stecher & G. W. Bohrnstedt (Eds.) *Class size reduction in California: The 1998-99 evaluation findings*. Sacramento, CA: California Department of Education.
- Stecher, B. M., McCaffrey, D.F., & Bugliari, D. (2003). The relationship between exposure to class size reduction and student achievement in California. *Education Policy Analysis Archives*, 11(40), 1-26.
- Woßmann, L., & West, M. (2006). Class-size Effects in School Systems around the World: Evidence from Between-grade Variation in TIMSS. *European Economic Review*, 50, 695–736.
- Word, E., Johnston, J., Pate-Bain, H., Fulton, B. D., Zaharias, J. B., Achilles, C. M., Lintz, M. N., Folgar, J., & Breda, C. (1990). *Student/Teacher Achievement Ratio (STAR): Tennessee's K 3 Class-size Study*. Nashville, TN: Tennessee State Department of Education.