Aust. Educ. Res. (2012) 39:59–73
DOI 10.1007/s13384-011-0047-3

Academic outcomes from between-class achievement grouping: the Australian primary context

Suzanne Macqueen

Published online: 17 December 2011

Abstract Grouping students by academic achievement level has been practised in a wide variety of forms and contexts for over a century. Despite a general consensus in the research that between-class achievement grouping provides no overall benefit for students, the practice has persisted in various guises. Between-class achievement grouping is common in high schools, and is also practised in a number of primary schools in various countries. While the affective outcomes of such practices have been investigated recently, academic outcomes at primary level have not been studied in recent decades. This paper examines the academic outcomes of between-class achievement grouping in literacy and numeracy classes in Australian primary schools. Results from standardised tests are compared between two groups of schools—one regroups students for these areas, and one maintains mixed-achievement classes. It is argued that the current regrouping practice closely resembles streaming and provides no apparent academic advantage for students.

Keywords Ability grouping · Primary education · Academic outcomes

Grouping students by achievement

Ongoing interest in the practice of grouping students by achievement level has ensured a steady stream of research on the topic over more than a century. Despite this, there remain gaps in our knowledge, and some areas have received little or no research attention. One such area is the effect of between-class achievement grouping on academic outcomes for primary students. Whilst academic outcomes are often cited as the reason for implementing such practices, no recent studies have investigated this area. What we know about academic achievement and between-class

S. Macqueen (✉)
School of Education, University of Newcastle, Newcastle, NSW 2208, Australia
e-mail: Suzanne.macqueen@newcastle.edu.au

grouping in primary schools is dated by a number of decades. More recent research has related to academic outcomes in secondary contexts or to affective outcomes in primary and secondary schools, and this research has come predominantly from the United Kingdom.

Grouping students according to achievement can take a number of forms, with varying terms used to describe particular models. Such grouping can occur between or within classes. Within-class grouping is generally employed for short activities during the school day, such as reading instruction, and is flexible. Such groups are organised by the class teacher, and can be reorganised at any time or disbanded altogether without impact on the class or school as a whole. Between-class grouping, by contrast, affects the structure of the school as a whole, with students of similar achievement levels grouped into one class. The term 'streaming' describes the practice of allocating students to homogeneous achievement classes based on some measure of each student's overall academic performance. Streaming became a popular method of allocating students to classes in larger schools after the introduction of standardised testing in the 1920s (Slavin 1987), although the practice can be traced back to the 1800s (Otto 1950 cited in Kulik and Kulik 1982, p. 415). Under streaming, a school with three classes in a particular grade would allocate those students considered more able to the 'A' class. Students performing not quite so well would constitute the 'B' class, whilst those considered least able would form the 'C' class. Regrouping has emerged seemingly as an alternative to streaming, and with similar effects (Wiliam and Bartholomew 2004; Hallam and Ireson 2006, 2007; MacIntyre and Ireson 2002), as will be outlined later in the paper. Under the regrouping strategy (known in the UK as 'setting') students are allocated to separate, homogeneous ability classes for literacy and numeracy lessons, with remaining subjects taught in mixed achievement classes.

The research on which this paper is based is part of a larger research study which investigated the practice of regrouping primary students by achievement for literacy and numeracy in Australian primary schools. Specific areas of investigation included the impact of regrouping on academic achievement, student attitudes towards school, and teaching practices. This paper focuses on the effects of regrouping on the academic achievement of primary students. Firstly, the literature on achievement grouping is reviewed. Results from studies conducted in primary schools are followed by those in secondary schools. A focus on studies related to academic outcomes is supported by a brief overview of other effects. Secondly, an outline of the method used in the current study is provided. Results are then presented and discussed in relation to overall academic outcomes and differences by achievement group level. Lastly, the conclusion presents the implications for practice. The aim of this article is to determine the effectiveness of current regrouping practices as employed in some Australian primary schools, in improving academic outcomes for students. Other aspects of the larger study will be presented in forthcoming papers.

The persistence of achievement grouping

In the 1960s and 1970s a number of studies discredited streaming (for example Barker Lunn 1970; Jackson 1964) on the basis of academic and social inequality.

Jackson's seminal work, based on a survey of 660 English primary schools, found that streaming provided limited advantages for some groups of students, whilst hindering the advancement of the vast majority. Other specific problems included inaccurate allocation of students to groups, inequitable allocation of teachers to groups, lack of student movement between groups and underestimation of the achievement of students in low streams. While streaming was largely discontinued in the UK (MacIntyre and Ireson 2002) and Australia, it has maintained a strong presence in many American school systems (Ansalone and Biafora 2004; Oakes 1985) where it is called 'tracking'.

Even in places where streaming has all but disappeared, other, more contemporary achievement grouping practices have remained, including between-class achievement grouping implemented for individual subject areas. In such a scheme, a student may be allocated to a homogeneous achievement class for English, and a different one for mathematics. This practice is known as 'setting' in the UK and is covered by the terms 'tracking' or 'regrouping' in the US. The term 'regrouping' will be used here, as that term was most often used by teachers in the current study. Whilst widely instituted in secondary schooling, the practice is also employed by a considerable number of primary schools, both in Australia and overseas. In the UK, increased use of the practice has been encouraged through a number of government initiatives associated with the National Literacy and Numeracy Strategies of 1998, for the expected benefit of student attainment (Hallam et al. 2003; MacIntyre and Ireson 2002). A resurgence of focus on achievement, at least at the governmental level (Hallam et al. 2004; Whitburn 2001; Wiliam and Bartholomew 2004) has been evident.

Research conducted in the first half of the twentieth century tended to focus on the achievement-related effects of the grouping practice, followed by an emphasis on educational equity, student self-concept and motivation in the latter half (Kulik and Kulik 1982). While it is not necessary to review this extensive literature for the purpose of this paper, it is appropriate to highlight the predominantly negative view of the practice concerning students' self-esteem (Oakes 1985), social outcomes (Hallinan and Sorensen 1985; Osterman 1998) and equity (MacIntyre and Ireson 2002; Oakes 1985; Wiliam and Bartholomew 2004). Achievement grouping has been found to cause changes in teacher expectations of students (Rosenthal and Jacobson 1968) and in teaching practices (Babad 1993), both with possible disadvantages for students. The only major benefit cited is that teachers find their job easier in achievement-based classes, due to the reduced range of student learning needs for which they need to cater (Ansalone and Biafora 2004).

While the impact of this regrouping practice on academic outcomes is often cited as the premise for its implementation, Slavin's (1987) oft-cited review of achievement grouping research in the form of a best-evidence synthesis found no overall academic benefit from achievement grouping, but did note some specific exceptions. He found positive effects for within-class grouping for mathematics, and for between-class grouping if implemented for reading and/or mathematics. These results came from 'Joplin plan' studies (Slavin 1987, p. 295) which took place in the 1950s and 1960s in the US, and which have similarities with the regrouping practice that is the focus of this paper. According to Slavin (1987) the

success of these was dependent on the following factors: groupings were reassessed regularly, and students remained in mixed achievement classes for most of the day. The latter condition may be difficult to attain if students are regrouped for two important curriculum areas, as was suggested by Ireson and Hallam (1999) in their more recent review of the literature.

In recent years a number of UK researchers have focused on regrouping in primary schools (for example, Davies et al. 2003; Hallam et al. 2003, 2004). These studies found that the processes involved in effectively implementing achievement grouping are 'complex and time consuming' (Davies et al. 2003, p. 57) and that many of the problems found with streaming also occur in setting (regrouping) (MacIntyre and Ireson 2002; Hallam and Ireson 2006, 2007). Primary students were aware of the grouping structures used (Hallam et al. 2004), but not all were happy with their group placements (MacIntyre and Ireson 2002). These recent primary school-based studies have reported little in regard to academic outcomes.

More studies on achievement grouping and academic outcomes can be found in relation to secondary students. A meta-analysis of 52 studies on achievement grouping in secondary schools was conducted by Kulik and Kulik (1982). They found no significant difference in academic achievement overall, but as with the primary-based research, there were exceptions. Extension programs for gifted and talented students were found to be beneficial, but those programs designed for struggling students were ineffective in improving academic outcomes. A second meta-analytic study in a range of school settings (including primary schools) had similar results. Streamed and regrouped classes were found to produce no academic benefits, while within-class and across-grade achievement grouping were slightly beneficial for attainment. Enriched and accelerated programs were found to be moderately beneficial to attainment for the students involved (Kulik and Kulik 1992). The researchers argue that these successes in achievement grouping arrangements depend on the degree of adjustment of the subject matter to suit students' achievement levels.

Findings from more recent secondary studies repeat those of Kulik and Kulik's (1982) earlier work, with high achieving students performing best in regrouped classes, and low achieving students making most progress when in mixed achievement classes. A study of over 600 students in UK secondary schools found improved achievement in mathematics for high achieving students, but no differences for English or science (Ireson et al. 2002). Achievement for similar students was affected by placement in different sets, according to Ireson et al. (2002), with those placed in low-achieving groups being disadvantaged. Similar results for mathematics students have come from studies in the UK (Wiliam and Bartholomew 2004), Belgium (Opdenakker and Van Damme 2001) and Israel (Linchevski and Kutscher 1998). Through a study with over 900 secondary mathematics students involving lesson observations, questionnaires and interviews, Boaler et al. (2000) found overall negative effects of setting for student learning. They linked the predominance of achievement grouping in UK schools with modest achievement standards compared to other nations. Both Boaler et al. (2000) and Burstein (1993, cited in Whitburn 2001, p. 425) suggested that setting may be the cause of low levels of achievement.

It appears that the effects of achievement grouping may differ for specific groups of students. Disadvantaged students have been found to suffer most as a result of achievement grouping. Babad's (1993) review of the literature on teachers' differential behaviour found that students in low achieving groups were likely to receive low quality instruction. Fewer high order thinking tasks were provided for students in low achievement classes, according to interviews with over 300 US secondary teachers (Raudenbush et al. 1993). Such differences may be linked to differential teacher expectations (Macqueen 2010). Rosenthal and Jacobson's (1968) classic experimental study demonstrated the effects of teacher expectations related to student potential. Students in an elementary school with streamed classes were pre-tested using a 'standard non-verbal intelligence test' (Rosenthal and Jacobson 1968, p. 175). Results showed a significant, positive difference for students randomly allocated to the 'high achievers' group, with such students from the middle stream gaining the most benefit. Results in numerous studies since that time have supported the earlier findings (Rubie-Davies et al. 2006). As an example, Wiliam and Bartholomew (2004) found that teachers had low expectations of students in low groups, whilst expectations of students in high groups were often too high.

Differences may also occur according to student gender. Gender differences have been found in the areas of overall academic achievement, engagement, behavioural problems, auditory processing problems, school-leaving age, and enjoyment of school (Rowe 2003). In most cases the overall concern has been that boys perform less well than girls in our education systems. Specific areas include overall academic achievement, engagement, behavioural problems, auditory processing problems, school-leaving age, and enjoyment of school (Rowe 2003). With regard to achievement grouping practices, Hallam and Ireson (2006) conducted a questionnaire with over 5,000 year nine students in England, and found that girls were more strongly in favour of regrouping than boys. Other research has shown girls to be disadvantaged by placement in the top group, where teachers may proceed at too fast a pace, with little concern for deep understanding (Boaler 1997b; Wiliam and Bartholomew 2004).

In summary, between-class achievement grouping continues in many primary schools, ostensibly to benefit academic achievement, but without research evidence of such outcomes. This paper seeks to provide quantitative evidence as to the effects of the practice in primary schools so that stakeholders may be better informed about the academic implications of the practice.

Regrouping study method

The research method in the larger study on which this paper draws was designed to investigate any differences arising from the use of the regrouping strategy in primary school contexts. Accordingly, two groups of primary schools (four in each group) from an urban area of NSW were included in the study for comparative purposes. (The education systems in the state of New South Wales cater for a population of over 600,000 primary students [NSW DET 2007]). One group of schools regrouped students (at least in years 3–6) by achievement for literacy and

numeracy sessions, and will be referred to as the 'regrouping schools'. The other group maintained mixed achievement classes for all subject areas and will be referred to as 'non-regrouping schools'.

Schools from similar socioeconomic areas were selected for inclusion in order to reduce variables. Initially, schools in the region known to practise regrouping were approached to participate. As those agreeing served areas with low socioeconomic status, similar schools were selected to complete the sample. All six State schools in the study were in the same 'Like School Group' (LSG) as determined by the Educational Measurement Directorate, NSW Department of Education and Training (DET). LSGs are determined according to factors including location, academic outcomes and socioeconomic status, using data collected through Basic Skills Test (BST) results, Socio Economic Indicators for Areas (SEIFA) and the Accessibility Remoteness Index of Australia (ARIA). The Catholic schools were not able to be assessed using this system, but they were located in close proximity to State schools in this category, which is Metro B, the second lowest category for metropolitan schools. It is noteworthy that all State schools involved were part of the Priority Schools Funding Program, and one of the non-regrouping schools was also included in the Priority Action Schools Program. Both programs were established by the NSW DET in order to support disadvantaged schools.

In NSW, Stage 3 covers years 5 and 6 of schooling, with students aged approximately 11 and 12 years, and in their last 2 years of primary school. Quantitative data from Stage 3 (years 5 and 6) students were collected, in the form of BST growth results. The BST, comprising a writing task and testing in reading and numeracy, was conducted annually in NSW state schools with primary students in years 3 and 5, prior to its replacement in 2008 by the National Assessment Program–Literacy and Numeracy (NAPLAN). Stage 3 students were selected because their BST growth data would be available. All students in Stage 3 were invited to participate, and numbers were only limited by consent received. Table 1 presents detail regarding student participants in the larger study. Not all data was available for each individual student, due to absences or school transfers. A low response rate is acknowledged. Variations in this rate may be attributed to school culture, but most of these schools commented that it was often difficult to have any forms returned by parents (from whom consent was needed), especially forms which

Table 1 Student participants

School	Year 5 girls	Year 5 boys	Year 6 girls	Year 6 boys	Total	Consent rate (%)
Regrouping A	4	2	0	0	6	11
Regrouping B	6	14	15	14	49	81
Regrouping C	4	2	5	6	17	24
Regrouping D	4	2	1	0	7	9
Non-regrouping E	8	10	3	6	27	18
Non-regrouping F	3	5	3	6	17	21
Non-regrouping G	2	5	4	7	18	30
Non-regrouping H	8	8	7	4	27	49

the community may find formal, lengthy or complex. The forms used in this study were designed to meet the ethics requirements of the relevant institutes, and may not have been well suited to the parents involved.

The two school groups used different strategies to determine the allocation of students to classes. Both school groups gathered assessment data from various school tasks, as well as teacher observation. After this point, methods differed. Non-regrouping schools formed classes of heterogeneous achievement (parallel classes), separating disruptive students. These classes stayed together, with the same teacher, for all lessons other than those presented in relief from face-to face sessions. Regrouping schools formed their home classes in a similar way to non-regrouping schools. Students were in home classes for administrative tasks and for lessons other than literacy and numeracy, such as science, social studies, health and art. For literacy and numeracy lessons, the students were allocated to a high, middle or low achieving class, as determined by student performance on a number of assessment tasks. In this way a student could be working with three different cohorts (including teachers) throughout the school day—one for literacy, a second for numeracy and a third for other subject areas covered in the home class. Regrouping was often not implemented in schools until Stage 2 (years 3 and 4). Exceptions were occasionally made in the placement of disruptive students, who could be placed in a higher group than their performance warranted, in an effort to reduce behavioural problems.

All regrouping schools manipulated numbers of students so that low achieving classes were smaller in size than middle and high classes. For three of these schools, this was achieved by utilising additional staff to create three regrouped classes from two home classes, using funding the schools received due to their status as disadvantaged schools. The fourth regrouping school, which did not receive the same funding, made the high and middle level groups larger in order to reduce the size of the low achieving classes.

This paper examines the BST growth results. In NSW primary schools, from 1989 until 2007, standardised tests were conducted in the areas of literacy, mathematics and writing for students in years 3 and 5 (replaced in 2008 by a national assessment program). The tests were compiled and marked by the NSW Department of Education (DET), and administered under strict guidelines. Schools, parents and students were provided with results which provided a snapshot of the students' performances in the tests and could be used to support future planning within the schools. The NSW DET collates 'growth results' for students who have completed both the year 3 and year 5 tests at the same school. Growth results are a type of change score, providing a value which indicates each student's growth in performance between the two sets of tests. Such data are often referred to as 'value added' (Hattie 2003). Growth results were chosen for analysis in this study, as these would most accurately describe the effects on achievement of the regrouping strategy, and they allowed for prior attainment. Whilst standardised tests such as these may not always accurately reflect students' learning (Alloway and Gilbert 1998; Wright et al. 1997), the results provide a useful tool for comparing basic achievement levels between schools operating under similar conditions. Whilst class or school-based assessment may have provided a more accurate picture of student

learning, it would have been very difficult to make accurate comparisons between classes and schools. Note that less data were available for writing assessment than for literacy and numeracy, due to a change in the reporting of results to some schools, which was beyond the control of the researcher.

BST growth results from students attending the two different school groups were compared using independent sample *t*-tests. Results for different group levels were compared using analysis of variance. Data were also recorded as to student grade, gender, regrouping class levels and home class. Cross-tabulation was conducted to demonstrate differences in regrouped class composition by gender and grade. Results between schools and classes in both school groups were compared using analysis of variance. Post hoc calculations were completed, and confirmed that non-significant results were not spurious. A significance level of $p < 0.05$ was applied in all analyses.

Regrouping effects

Academic outcomes

Comparison of BST growth data related to literacy, mathematics and writing, using independent sample *t*-tests, shows no significant difference between regrouping and non-regrouping school groups as shown in Table 2. Significance levels are $p = 0.279$ for literacy growth, $p = 0.497$ for numeracy growth and $p = 0.727$ for writing growth. Growth results in writing are very similar between the two school groups, whilst non-significant differences favour regrouping schools for literacy and non-regrouping schools for mathematics. No significant differences were found between individual schools or classes.

The results demonstrate that regrouping provided no benefits in academic achievement and are consistent with overall results on achievement grouping found by Slavin (1987) and Kulik and Kulik (1982, 1992). The positive results from Joplin Plan strategies in the middle of last century described by Slavin (1987) were not replicated in the current study. It may be the case that certain conditions outlined as

Table 2 Students' growth in academic achievement by grouping structure

Structure	*N*	Mean	SD	Sig. (2-tailed)
BST growth in literacy				
Regrouping	50	7.29	3.33	0.279
Non-regrouping	68	6.50	4.49	
BST growth in mathematics				
Regrouping	51	6.75	5.82	0.497
Non-regrouping	69	7.44	5.25	
BST growth in writing				
Regrouping	29	5.38	3.60	0.727
Non-regrouping	47	5.83	6.35	

being crucial for success were not mirrored in this study, such as accurate allocation of students to groups, fluid movement of students between groups, as indicated by ongoing evaluation, or the maintenance of heterogeneous classes for most of the day. Indeed, students in regrouping schools in this study were in heterogeneous classes for less than half the school day.

Tailoring the curriculum to suit student needs closely, a factor Kulik and Kulik (1992) cited as producing benefits in gifted and talented classes, may be difficult in the case of the larger high achievement groups in this study—these often contained more than 40% of the Stage cohort. It may be that regrouped classes do not cater for gifted and talented students in particular, with some students being allocated to the high groups purely to keep low groups small. Also, students may vary in achievement level within a subject area. For example, a student may excel in one strand of mathematics such as number but struggle with measurement, and such differences may not be considered by teachers of achievement based classes.

Differences by gender

Given recent concerns about the academic performance of boys (Rowe 2003), some differences by gender may have been expected in academic achievement overall. Initial analysis of the BST growth data using independent sample *t*-tests in relation to gender alone, showed no overall significant difference in means between results achieved by boys and girls, as shown in Table 3. That is, there was no significant difference in the academic performance of boys and girls. Mean growth for boys was slightly higher in literacy, but lower in mathematics and writing. The lack of difference shown here may be due to the use of growth data, which accounts for prior attainment, with these results covering a 2 year period only.

When grouping structure was incorporated as a variable, analysis of the BST growth data also showed no significant difference by gender (as shown in Tables 4, 5). The greatest non-significant difference was for boys in regrouping schools (mean = 5.69), who demonstrated considerably less growth in mathematics than boys in non-regrouping schools (mean = 8.00).

Table 3 Students' growth in academic achievement by gender

Gender	*N*	Mean	SD	Sig. (2-tailed)
BST growth in literacy				
Boys	62	7.14	3.80	0.387
Girls	56	6.49	4.31	
BST growth in mathematics				
Boys	63	6.97	5.49	0.708
Girls	57	7.35	5.53	
BST growth in writing				
Boys	41	5.37	6.21	0.757
Girls	35	5.76	4.47	

Table 4 Boys' growth in academic achievement by grouping structure

Structure	*N*	Mean	SD	Sig. (2-tailed)
BST growth in literacy				
Regrouping	27	7.40	2.88	0.630
Non-regrouping	35	6.95	4.41	
BST growth in mathematics				
Regrouping	28	5.69	4.87	0.098
Non-regrouping	35	8.00	5.81	
BST growth in writing				
Regrouping	15	6.05	2.56	0.606
Non-regrouping	26	4.99	7.59	

Table 5 Girls' growth in academic achievement by grouping structure

Structure	*N*	Mean	SD	Sig. (2-tailed)
BST growth in literacy				
Regrouping	23	7.16	3.86	0.341
Non-regrouping	33	6.03	4.60	
BST growth in mathematics				
Regrouping	23	8.05	6.69	0.438
Non-regrouping	34	6.88	4.64	
BST growth in writing				
Regrouping	14	5.61	4.55	0.869
Non-regrouping	21	5.87	4.52	

Examination of data from girls alone, shown in Table 5, likewise shows no significant difference in academic performance between the two school groups. Non-significant differences show higher mean results for girls in regrouping schools for both literacy and mathematics, whilst writing results were very similar.

No difference by gender can be interpreted from these results. There were no significant differences between results from regrouping and non-regrouping schools for gender in literacy, numeracy or writing. Although boys achieved less in regrouped mathematics classes and girls achieved more in regrouped classes for both literacy and mathematics, the differences were not statistically significant. This difference is consistent with general trends for gender, where boys have been seen to perform less well academically than girls (Rowe 2003), but may also be affected by boys' placement in groups.

Boys were over-represented in the low achieving groups for both literacy ($\chi^2 = 6.64$, df $= 2$, $p = 0.036$) and mathematics ($\chi^2 = 12.27$, df $= 2$, $p = 0.002$) (see Tables 6, 7), whilst there were similar percentages of each gender in middle groups; girls dominated the high groups for both subjects.

These differences in student allocation to groups by gender for both literacy and numeracy groups were statistically significant, and may represent a disadvantage to boys in the regrouping strategy.

Table 6 Literacy group level placement by gender

Gender	Low (%)	Middle (%)	High (%)	Total number of students
Boys	24	34	42	38
Girls	3	38	59	34

$\chi^2 = 6.64$, df $= 2$, $p = 0.036$

Table 7 Mathematics group level placement by gender

Gender	Low (%)	Middle (%)	High (%)	Total number of students
Boys	21	39.5	39.5	38
Girls	6	34	60	34

$\chi^2 = 12.27$, df $= 2$, $p = 0.002$

Differences by achievement group level

Allocation to regrouped classes

The collation of student group placement information showed that the majority of students from regrouping schools were effectively in a streamed situation. Of the 78 students from regrouping schools who were surveyed, the majority (56 students or approximately 70%) were in the same achievement group level for English and for mathematics instruction. Seven students were in a higher group for English than mathematics, eight were in a higher group for mathematics. In only one identified case was a student placed in achievement groups which were more than a level apart (that is, the student was in a low group for English and a high group for mathematics). Seven students had incomplete data in this area, meaning that the percentage of students in the same group level for both literacy and mathematics could be greater than 70%. It is important to recognise this, as these students are effectively in a streamed class for most of the school day—that system so thoroughly criticised by Jackson (1964) and others (Ansalone and Biafora 2004; Kulik and Kulik 1982; Oakes 1985; Slavin 1987). This effect has negative implications for both cognitive and affective results—the latter to be discussed in further papers (Macqueen, under review).

The achievement level of the group to which a student belonged did not have a significant effect, according to results obtained in this study. Four achievement group levels were compared. A student's group level could be low, middle or high in regrouping schools or mixed in non-regrouping schools. Analysis of variance applied to BST growth data showed no significant difference in results for either mathematics group level or literacy group level. The results generated by this analysis are shown in Tables 8 and 9.

Although this is not significantly different, low achieving mathematics students produced a lower mean growth in mathematics achievement (but not for literacy or writing) than did other groups, as shown in Table 8.

Table 8 Students' growth in academic achievement by mathematics group level

Group level	*N*	Mean	SD	*F*	Sig.
BST growth in literacy					
Low	6	7.67	4.07	0.78	0.538
Middle	15	6.29	2.59		
High	24	7.76	3.61		
Mixed	68	6.50	4.49		
BST growth in mathematics					
Low	6	3.35	5.51	1.39	0.248
Middle	16	7.13	4.15		
High	24	8.40	6.57		
Mixed	69	7.44	5.25		
BST growth in writing					
Low	3	6.40	2.04	0.07	0.976
Middle	12	5.60	3.20		
High	8	6.20	5.34		
Mixed	47	5.37	6.35		

Table 9 Students' growth in academic achievement by literacy group level

Group level	*N*	Mean	SD	*F*	Sig.
BST growth in literacy					
Low	5	4.86	1.68	0.94	0.424
Middle	15	7.71	2.99		
High	25	7.47	3.70		
Mixed	67	6.53	4.52		
BST growth in mathematics					
Low	5	7.58	4.30	0.06	0.981
Middle	16	6.87	3.97		
High	25	7.52	7.11		
Mixed	68	7.49	5.28		
BST growth in writing					
Low	3	5.60	1.22	0.14	0.935
Middle	9	5.22	2.82		
High	11	6.57	4.97		
Mixed	46	5.36	6.42		

Likewise, the mean growth for low achieving literacy students was lower for literacy than other groups, but the difference was not significantly different.

Whilst the differences shown here are not statistically significant, mean growth was lower in both mathematics and literacy, for students in low groups for each subject. Further studies with a larger sample size, particularly of low achieving

students, would be useful, as the sample was limited in this study. This was due, in part, to the choice made by regrouping schools, to make low achievement classes smaller in size. These results suggest that low achieving primary students may be disadvantaged by achievement-based classes, as suggested by much of the previous research in secondary schools (Ireson et al. 2002; Wiliam and Bartholomew 2004). This may or may not be related to teacher expectation, as discussed earlier. It is interesting to note that these results occurred despite low achievement classes being smaller in size than other groups; this also brings into question the practice of keeping low achieving classes smaller. Research has suggested that small class size can benefit academic outcomes, but some claim that the benefits only exist with classes of less than 20 (Blatchford and Mortimore 1994; Finn et al. 2003), and for students in the early years of schooling (Blatchford and Mortimore 1994). A study by Wright et al. (1997) found only minor effects for students in Grades 3–5. This is mirrored by the lack of result shown with the Stage 3 students studied here. These students may benefit more from mixed achievement classes where there are opportunities for peer tutoring, role modelling and exposure to a wider range of teaching strategies (Boaler 1997a).

Conclusion

This study provides evidence in line with previous research, that regrouping primary students by achievement for literacy and numeracy classes is ineffective in producing academic gains, and also reveals that for many students it effectively creates a streamed class situation. Care must be taken when analysing the data used in this paper, due to the low response rate. However, evidence from the larger study (including teacher interviews and observations as well as completion of the Quality of School Life survey by students) provides no contrasting insight; rather, it supports the hypothesis that regrouping affects teacher attitudes and practices, as well as social aspects of schooling (Macqueen 2009), with undesirable outcomes.

When regrouped for literacy and numeracy, the majority of primary students are effectively streamed, yet principals and teachers seem largely unaware that this is the case, and few would recommend a return to that practice. Schools that continue to employ regrouping strategies must work to overcome the problems of these strategies which, whilst not being insurmountable, have indeed persisted since the days of widespread streaming (MacIntyre and Ireson 2002; Hallam and Ireson 2006, 2007). Strategies that meet the needs of teachers and students and that do not rely on between-class achievement grouping should be investigated. Preservice and inservice professional development related to effective mixed-achievement teaching would seem an obvious place to start (Ansalone and Biafora 2004; Linchevski and Kutscher 1998; Wiliam and Bartholomew 2004). In the case of schools investing extra resources in the regrouping (such as the extra staff used by the State schools to lower class size), such resources may be better allocated, perhaps through professional development.

It is concerning that between-class achievement grouping practices such as those described in this study, persist in our schools, despite the lack of supporting

evidence for them, so it may be that other explanations for their continuation need to be sought.

Acknowledgments The author would like to acknowledge the feedback from members of the School of Education writing groups at the University of Newcastle on an earlier draft of this paper, and from the anonymous peer reviewers, which helped to get this work published.

References

Alloway, N., & Gilbert, P. (1998). Reading literacy test data: Benchmarking success. *The Australian Journal of Language and Literacy, 21*(3), 249–261.

Ansalone, G., & Biafora, F. (2004). Elementary school teachers' perceptions and attitudes to the educational structure of tracking. *Education, 125*(2), 249–258.

Babad, E. (1993). Teachers' differential behaviour. *Educational Psychology Review, 5*(4), 347–376.

Barker Lunn, J. C. (1970). *Streaming in the primary school: A longitudinal study of children in streamed and non-streamed junior schools*. Slough: National Foundation for Educational Research.

Blatchford, P., & Mortimore, P. (1994). The issue of class size for young children in schools: What can we learn from research? *Oxford Review of Education, 20*(4), 411–428.

Boaler, J. (1997a). Setting, streaming and mixed ability teaching. In J. Dillon & M. Maguire (Eds.), *Issues in secondary teaching. Becoming a teacher*. Buckingham: Open University Press.

Boaler, J. (1997b). *Experiencing school mathematics: Teaching styles, sex and setting*. Buckingham: Open University Press.

Boaler, J., Wiliam, D., & Brown, M. (2000). Students' experiences of ability grouping—disaffection, polarisation and the construction of failure. *British Educational Research Journal, 26*(5), 631–648.

Davies, J., Hallam, S., & Ireson, J. (2003). Ability groupings in the primary school: Issues arising from practice. *Research Papers in Education, 18*(1), 45–60.

Finn, J. D., Pannozzo, G. M., & Achilles, C. M. (2003). The 'Why's' of class size: Student behaviour in small classes. *Review of Educational Research, 73*(3), 321–368.

Hallam, S., & Ireson, J. (2006). Secondary school pupils' preferences for different types of structured grouping practices. *British Educational Research Journal, 32*(4), 583–599.

Hallam, S., & Ireson, J. (2007). Secondary school pupils' satisfaction with their ability grouping placements. *British Educational Research Journal, 33*(1), 27–45.

Hallam, S., Ireson, J., & Davies, J. (2004). Primary Pupils' experiences of different types of grouping in school. *British Educational Research Journal, 30*(4), 515–533.

Hallam, S., Ireson, J., Lister, V., Andon Chaudhury, I., & Davies, J. (2003). Ability grouping practices in the primary school: A survey. *Educational Studies, 29*(1), 69–83.

Hallinan, M. T., & Sorensen, A. B. (1985). Ability grouping and student friendships. *American Educational Research Journal, 22*(4), 485–499.

Hattie, J. (2003). *Teachers make a difference: What is the research evidence*? Paper presented at the Australian Council for Educational Research Conference, Melbourne.

Ireson, J., & Hallam, S. (1999). Raising standards: Is ability grouping the answer? *Oxford Review of Education, 25*(3), 343–358.

Ireson, J., Hallam, S., Hack, S., Clark, H., & Plewis, I. (2002). Ability grouping in english secondary schools: Effects on attainment in english, mathematics and science. *Educational Research and Evaluation, 8*(3), 299–318.

Jackson, B. (1964). *Streaming, an education system in miniature*. London: Routledge & Kegan Paul Ltd.

Kulik, J. A., & Kulik, C.-L. C. (1982). Effects of ability grouping on secondary school students: A meta-analysis of evaluation findings. *American Educational Research Journal, 19*(3), 415–428.

Kulik, J. A., & Kulik, C.-L. C. (1992). Meta-analytic findings on grouping programs. *Gifted Child Quarterly, 36*(2), 73–77.

Linchevski, L., & Kutscher, B. (1998). Tell me with whom you're learning, and I'll tell you how much you've learned. *Journal for Research in Mathematics Education, 29*(5), 533–554.

MacIntyre, H., & Ireson, J. (2002). Within-class ability grouping: Placement of pupils in groups and self-concept. *British Educational Research Journal, 28*(2), 249–263.

Macqueen, S. (2009). *Grouping Primary students by achievement for literacy and numeracy instruction: Who wins*? Paper presented at National Conference for Teachers of English and Literacy, Hobart.

Macqueen, S. (2010). Primary teacher attitudes in achievement-based literacy classes. *Issues in Educational Research*, 20(2), 118–136.
NSW DET (2007). Planning and innovation. *Statistical Bulletin: Schools and Students in New South Wales, 2006*. https://www.det.nsw.edu.au/media/downloads/reports. Accessed September, 17, 2008.
Oakes, J. (1985). *Keeping track: How schools structure inequality*. New Haven, CT: Yale University Press.
Opdenakker, M., & Van Damme, J. (2001). Relationship between school composition and characteristics of school process and their effect on mathematics achievement. *British Educational Research Journal, 27*(4), 407–432.
Osterman, K. F. (1998). *Student community within the school context: A research synthesis*. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego.
Raudenbush, S. W., Rowan, B., & Cheong, Y. F. (1993). Higher order instructional goals in secondary schools: Class, teacher and school influence. *American Educational Research Journal, 30*(3), 523–553.
Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. New York: Holt, Rinehart and Winston, Inc.
Rowe, K. J. (2003). *The importance of teacher quality as a key determinant of students' experiences and outcomes of schooling*. Background paper to keynote address presented at the Australian Council of Educational Research Conference, Melbourne, 19–21.
Rubie-Davies, C., Hattie, J., & Hamilton, R. (2006). Expecting the best for students: Teacher expectations and academic outcomes. *British Journal of Educational Psychology, 76*(3), 429–444.
Slavin, R. E. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. *Review of Educational Research, 57*(3), 293–336.
Whitburn, J. (2001). Effective classroom organisation in primary schools: Mathematics. *Oxford Review of Education, 27*(3), 411–428.
Wiliam, D., & Bartholomew, H. (2004). It's not which school but which set you're in that matters: The influence of ability grouping practices on student progress in mathematics. *British Educational Research Journal, 30*(2), 279–293.
Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education, 11*, 57–67.

Author Biography

Suzanne Macqueen is a lecturer in the School of Education at the University of Newcastle, teaching courses related to literacy, quality teaching and professional preparation. She completed a Master of Education (Research) on the topic of between-class achievement grouping for literacy and numeracy classes in primary schools. Her PhD research examines the experiences of non-traditional students in teacher education programs.