MAC08085

Between-class achievement grouping for literacy and numeracy: academic outcomes for primary students

AARE 2008 Conference in Brisbane QLD
30 November-4th December 2008
University of Technology
Changing Climates: Education for Sustainable Futures

Suzanne Macqueen, University of Newcastle, Australia
Suzanne.macqueen@newcastle.edu.au
Theme: Primary Education

Abstract

Achievement grouping has been practised in a number of forms and contexts for over a century, and has been the subject of copious amounts of research. 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. One form is that which occurs for individual subject areas. This is quite common in high schools, particularly for mathematics, and is also practised in a number of primary schools in New South Wales and overseas. Recent research in this field involving primary school students and teachers has investigated the affective outcomes of such practices, but academic outcomes at the primary level have not been studied in recent decades. This paper examines the academic outcomes of between-class achievement grouping in primary literacy and numeracy classes. Results from Basic Skills tests are compared between two groups of schools – one which regroups students for these areas, and one where students remain in a mixed-achievement class for all subjects. The conclusion reached is that the current regrouping practice provides no academic advantage for students.

What we know about between-class grouping

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 areas which have received little or no attention from researchers. 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 research has investigated this area. What we know about academic achievement and between-class 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 schools, predominantly in the United Kingdom (UK). This paper addresses the gap in the literature by presenting the findings of a contemporary study on between-class achievement grouping in Australian primary schools. The focus here will be on student academic outcomes in the

areas of literacy and numeracy, whilst other aspects of the larger study will be presented elsewhere.

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 designated by the class teacher, and can be reorganised at any time or disbanded altogether without impact on the class 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 homogenous 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 & 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 least able would form the "C" class. A number of studies discredited streaming in the 1960s and 1970s (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. Another difficulty was that a student may have differing capabilities in different subject areas, such as being a high achiever in mathematics but poor in literacy. While streaming was largely discontinued in the United Kingdom (UK) (MacIntyre & Ireson, 2002) and other countries, including Australia, it has maintained a strong presence in many American school systems (Ansalone & Biafora, 2004; Oakes, 1985) where it is called "tracking".

Even in places where streaming has all but disappeared, other achievement grouping practices have remained, and research in that broad area has been plentiful. One common contemporary use of achievement grouping is between-class achievement grouping implemented for individual subject areas. That is, 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 covered by the terms "tracking" and "regrouping" in the US. 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 since the 1990s by a number of governmental initiatives associated with the literacy and numeracy hours for the expected benefit of student attainment

(Hallam, Ireson, Lister, Andon Chaudhury & Davies, 2003; MacIntyre & Ireson, 2002). This illuminates a resurgence of focus on achievement, at least at the governmental level (Hallam, et al., 2004; Whitburn, 2001; Wiliam & Bartholomew, 2004).

The focus of research into achievement grouping has changed over the decades. That conducted in the first half of the twentieth century tended to focus on the achievement-related effects of the practice, followed by an emphasis on equality of opportunity, student self-concept and motivation in the latter half (Kulik & Kulik, 1982). Whilst it is not necessary to review all of that literature for the purpose of this paper, a summary may be useful in providing context. Overall, researchers have voiced concern about possible negative affects of achievement grouping for students' self-esteem (Oakes, 1985), social outcomes (Hallinan & Sorensen, 1985; Osterman, 1998) and equity (MacIntyre & Ireson, 2002; Oakes, 1985; Wiliam & Bartholomew, 2004). Additionally, achievement grouping has been found to cause changes in teacher expectations of students and in teaching practices, both with possible disadvantages for students. Teachers find their job easier in achievement based classes due to the reduced range of student learning needs (Ansalone & Biafora, 2004).

The impact of this regrouping practice on academic outcomes is of particular interest, as this is often claimed as the premise for its implementation. Slavin's (1987) oft cited review of achievement grouping research in the form of a best-evidence synthesis related specifically to primary level education and included studies on both between- and within-class achievement grouping. He found no overall academic benefit from achievement grouping, but did note some specific grouping techniques which were exceptions. He found positive effects from within-class grouping for mathematics, and from between-class grouping if done for reading and/or mathematics. These particular results came from "Joplin plan" studies (Slavin, 1987, p. 295) which took place in the 1950s and 1960s in the United States, and which have similarities with the regrouping practice on which this paper will focus. These 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 has focused on regrouping in primary schools (for example, Davies, Hallam, & Ireson, 2003; Hallam, Ireson, & Davies, 2004; Hallam, et al., 2003). Results from these studies state 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 & Ireson, 2002; Hallam & Ireson, 2006, 2007). Achievement grouping was found to be most often employed for mathematics and English, and occurred with increasing frequency as students progressed through school (Hallam, et al., 2003). Primary students were aware of the grouping structures used (Hallam, et al., 2004), but not all were happy with

their group placements (MacIntyre & Ireson, 2002). These recent primary school based studies have reported little as to academic outcomes.

More studies on academic outcomes and achievement grouping can be found in relation to secondary students. A meta-analysis of 52 studies conducted by Kulik and Kulik (1982) on achievement grouping in secondary schools similarly 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. These researchers conducted a second meta-analytic study related to a range of school settings (including primary schools) with similar results. Streamed and setted classes were found to produce no academic benefits, within-class and across-grade (most similar to regrouping) achievement grouping were slightly beneficial for attainment, while enriched and accelerated programs were moderately beneficial to attainment (Kulik & Kulik, 1992). The researchers suggested that the success of some achievement grouping arrangements lay in the degree of adjustment of the subject matter to suit students' achievement levels.

Generally, findings from more recent secondary studies have reflected those from Kulik & Kulik's (1982) earlier study, with high achieving students performing best in regrouped classes, and low achieving students making most progress when in mixed achievement classes. A study using data from over 600 students in UK secondary schools found improved achievement in mathematics for high achieving students, but no differences for English or science (Ireson, Hallam, Hack, Clark & Plewis, 2002). Achievement for similar students was affected by placement in different sets, according to Ireson, Hallam et al. (2002), with those placed in low-achieving groups disadvantaged. Similar results for mathematics students have come from studies in the UK (Wiliam & Bartholomew, 2004), Belgium (Opendakker & Van Damme, 2001) and Israel (Linchevski & Kutscher, 1998). Through a study with over 900 secondary mathematics students involving lesson observations, questionnaires and interviews, Boaler, Wiliam and Brown (2000) found that setting affected student learning negatively. 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 is possible that effects 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, Rowan & Cheong, 1993). Such differences may be linked to differential teacher expectations.

Rosenthal and Jacobson's (1968) classic experimental study demonstrated the effects of teacher expectations related to student potential, and the results have been confirmed in numerous studies since that time (Rubie-Davies, Hattie & Hamilton, 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.

Gender is another area where differences may occur. 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 fare less well in our education systems. With regard to achievement grouping practices, Hallam and Ireson (2006) conducted a questionnaire with over 5000 Year 9 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, 1997; Wiliam & Bartholomew, 2004).

The study on which this paper is based was conducted as part of a research higher degree, and aimed to investigate the practice of regrouping primary students by achievement for literacy and numeracy in Australian primary schools. Specific areas of investigation included the impact of the practice on academic achievement, student attitudes towards school, and teaching practices. This paper will present results related to the effects of regrouping on academic achievement, with the aim of determining the effectiveness of current regrouping practices as employed in some Australian primary schools in improving academic outcomes for their students.

Regrouping Study Method

Two groups of primary schools (four in each group) were included in the study. 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 "regrouping schools". In these schools, students may belong to three separate classes: one for literacy, one for mathematics and another for remaining subject areas. The other group maintained mixed achievement classes for all subject areas and will be referred to as "non-regrouping schools". All regrouping schools manipulated numbers of students so that low achieving classes were smaller in size than middle and high classes. Schools from similar socioeconomic areas were selected for inclusion in order to reduce variables;
as initial schools agreeing to participate in the study served areas with low socioeconomic status, similar schools were selected to complete the sample. Interviews about the regrouping strategy and its effects on teachers, students and teaching practices were conducted with principals and teachers, and some classroom observations were completed. Quantitative data from Stage

3 (Years 5 and 6) students were also collected, in the form of Basic Skills Test (BST) growth results and Quality of School Life survey responses.

In this paper, BST growth results are examined. 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 (and was replaced in 2008 by a national assessment program). The tests were compiled and marked by the NSW Department of Education (DET), and were 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. Where Year 5 students completed both the Year 3 and Year 5 tests at the same school, growth results were also provided. These growth results provided students with a value which indicated each student's growth in performance between the two sets of tests. Such data are often referred to as "value added" (Hattie, 2003). The growth results were chosen for analysis in this study, as it was thought that these would most accurately describe the effects on achievement of the regrouping strategy, as this strategy was often not implemented in schools until Stage 2 (Years 3 and 4). An added advantage of using growth results is that they account for prior attainment, which is known to affect student academic outcomes (Hattie, 2003). In cases where students were absent for any of the BST tests, or had completed them at different schools, these data were not able to be obtained. Whilst there is a widely held belief that standardised tests such as these may not accurately reflect students' learning (Alloway & Gilbert, 1998; Wright, Horn & Sanders, 1997), these results provide a useful tool for comparing basic achievement levels between schools operating under similar conditions.

BST growth results from the two different school groups were compared using independent sample *t*-tests. The effects of gender were also measured in this way. Results for different group levels were compared using analysis of variance.

Data were recorded as to student grade, gender, regrouping class levels and home class. Crosstabulation 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.

Regrouping Results and Reflections

Academic outcomes

Comparison of BST growth data related to literacy, mathematics and writing using independent sample *t*-tests showed no significant difference between regrouping and non-regrouping school groups at $p < .05$ as shown in Table 1. 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.

Table 1. Students' growth in academic achievement by grouping structure

	Structure	N	Mean	Std dev.	Sig. (2-tailed)
BST growth in literacy	Regrouping Non-regrouping	50 68	7.29 6.50	3.33 4.49	.279
BST growth in mathematics	Regrouping Non-regrouping	51 69	6.75 7.44	5.82 5.25	.497
BST growth in writing	Regrouping Non-regrouping	29 47	5.38 5.83	3.60 6.35	.727

The results demonstrate that regrouping provides no benefits in academic achievement. This is consistent with overall results on achievement grouping found by Slavin (1987) and Kulik and Kulik (1982, 1992), but does not support some of their specific findings. The results from Joplin Plan strategies described by Slavin were not replicated in the current study. It may be the case that certain conditions outlined by Slavin (1987) as being crucial for success were not mirrored in this study, such as accurate allocation of student 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. Another possible reason that academic benefits were not seen is that the adequate differentiation of the curriculum to suit student skills and needs, as advocated by Kulik and Kulik (1992), may not have occurred. It would be especially difficult to tailor curriculum to closely suit student needs in the case of the larger high achievement groups in this study, which often contained more than 40 per cent of the Stage cohort. Also, students may vary in achievement level within a subject area. For example, a student may excel with one strand of mathematics such as number but struggle with data, and such differences may not be considered by teachers with achievement based classes.

Differences by gender

Initial analysis of the BST growth data using independent sample *t*-tests in relation to gender alone showed no overall significant difference in means at $p < .05$ between results achieved by boys and girls, as shown in Table 2. When grouping structure was incorporated as a variable, analysis of the BST growth data also showed no significant difference at $p < .05$ by gender (as shown in Tables 3 and 4), although boys in regrouping schools demonstrated considerably less growth in mathematics.

Table 2 Students' growth in academic achievement by gender

	Gender	N	Mean	Std dev.	Sig. (2-tailed)
BST growth in literacy	Boys Girls	62 56	7.14 6.49	3.80 4.31	.387
BST growth in mathematics	Boys Girls	63 57	6.97 7.35	5.49 5.53	.708

BST growth in writing	Boys Girls	41 35	5.37 5.76	6.21 4.47	.757

Table 3 Boys' growth in academic achievement by grouping structure

	Structure	N	Mean	Std dev.	Sig. (2-tailed)
BST growth in literacy	Regrouping Non-regrouping	27 35	7.40 6.95	2.88 4.41	.630
BST growth in mathematics	Regrouping Non-regrouping	28 35	5.69 8.00	4.87 5.81	.098
BST growth in writing	Regrouping Non-regrouping	15 26	6.05 4.99	2.56 7.59	.606

Table 4 Girls' growth in academic achievement by grouping structure

	Structure	N	Mean	Std dev.	Sig. (2-tailed)
BST growth in literacy	Regrouping Non-regrouping	23 33	7.16 6.03	3.86 4.60	.341
BST growth in mathematics	Regrouping Non-regrouping	23 34	8.05 6.88	6.69 4.64	.438
BST growth in writing	Regrouping Non-regrouping	14 21	5.61 5.87	4.55 4.52	.869

No difference by gender can be interpreted from these results. Although boys achieved less in regrouped mathematics classes, the differences were not statistically significant. This difference is consistent with general trends (Rowe, 2003), but may relate to boys' placement in groups. As seen in Tables 5 and 6, boys were over-represented in the low achieving groups for both literacy and mathematics, whilst there were similar percentages of each gender in middle groups, and girls dominated the high groups for both subjects. Differences in student allocation to groups by gender for both literacy and numeracy groups were statistically significant.

Table 5 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-sq=6.64, df=2, p=.036)

Table 6 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-sq=12.27, df=2, p=.002)

Differences by achievement level

In the study, there were four possible levels for a student to belong to for literacy and numeracy classes. These were low, middle or high for students in regrouping schools, and mixed for students in non-regrouping schools. 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 as for mathematics instruction. Seven students were in a higher group for English than mathematics and eight 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 per cent. It is important to recognise this, as these students are effectively in a streamed class for most of the school day, as was previously suggested by Ireson and Hallam (1999).

Analysis of variance applied to BST growth data showed no significant difference in results for either mathematics group level or literacy group level. Although not significantly different, low achieving mathematics students produced a lower mean growth in mathematics achievement than did other groups. Likewise, the mean growth for low achieving literacy students was lower for literacy than other groups. The results generated by this analysis are shown in Tables 7 and 8.

Table 7 Students' growth in academic achievement by mathematics group level

	Group level	N	Mean	Std Dev	F	Sig.
BST growth in literacy	Low	6	7.67	4.07	.78	.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	.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	.07	.976
	Middle	12	5.60	3.20		
	High	8	6.20	5.34		
	Mixed	47	5.37	6.35		

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

	Group level	N	Mean	Std dev.	F	Sig.
BST growth in literacy	Low	5	4.86	1.68	.94	.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	.06	.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	.14	.935
	Middle	9	5.22	2.82		

	High Mixed	11 46	6.57 5.36	4.97 6.42		

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, Hallam, et al., 2002; Wiliam & Batholomew, 2004). It is interesting to note that these results occurred despite low achievement classes being smaller in size than other groups, which also brings that strategy into question.

Conclusion

This study provides evidence that, in line with the previous research, regrouping primary students by achievement for literacy and numeracy classes is ineffective in producing academic gains. School stakeholders' justifications for the practice are not supported. It is concerning that between-class achievement grouping practices such as this persist in our schools despite the lack of supporting evidence, and it must be considered that there are other reasons for the continuation. It may be that principals like to be seen as taking action or that teachers prefer this system because it makes their job easier (Ansalone & Biafora, 2004; Hallam & Ireson, 2003, 2005). Whilst making the teacher's job easier is not without value, given that such improvements may have flow on benefits for students also, it is unwise to pursue this as a goal by ignoring lessons from research.

When regrouped for literacy and numeracy, the majority of primary students are effectively streamed, yet principals and teachers seem largely ignorant that this is the case, and few of them would recommend a return to that practice. Schools which continue to employ regrouping strategies must work to overcome the problems of these strategies which, whilst not insurmountable, have indeed remained since the days of widespread streaming (MacIntyre & Ireson, 2002; Hallam & Ireson, 2006, 2007). It is preferable that strategies for meeting the needs of teachers and students which do not rely on between-class achievement grouping should be implemented. Preservice and inservice professional development related to effective mixed-achievement teaching would seem an obvious place to start (Ansalone & Biafora, 2004; Wiliam & Bartholomew, 2004).

Educators, along with all who desire equity in our society, must promote socially just goals and practices. Jackson (1964) suggested that society's values had led to streaming. It would seem that these values have remained largely unchanged, and have resurfaced in a new format as regrouping or setting. It may not be until these values have changed that such practices are consigned to the past.

References

Alloway, N. & Gilbert, P. (1998). Reading literacy test data: Benchmarking success.

The Australian Journal of Language and Literacy, 21(3), pp. 249-261.

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

Babad, E. (1993). Teachers' Differential Behaviour. *Educational Psychology Review,* 5(4), pp. 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.

Boaler, J. (1997). *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), pp. 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), pp. 45-60.

Hallam, S. & Ireson, J. (2003). Secondary school teachers' attitudes towards and beliefs about ability grouping. *British Journal of Educational Psychology*, 73(3), pp. 343-356.

Hallam, S. & Ireson, J. (2005). Secondary school teachers' pedagogic practices when teaching mixed and structured ability classes. *Research Papers in Education,* 20(1), pp. 3-24.

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

Hallam, S. & Ireson, J. (2007). Secondary school pupils' satisfaction with their ability grouping placements. *British Educational Research Journal,* 33 (1), pp. 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), pp. 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), pp. 69-83.

Hallinan, M. T. & Sorensen, A. B. (1985). Ability Grouping and Student Friendships. *American Educational Research Journal,* 22(4), pp. 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., 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), pp. 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), pp. 415-428.

Kulik, J. A. & Kulik, C.-L. C. (1992). Meta-analytic Findings on Grouping Programs. *Gifted Child Quarterly,* 36(2), pp. 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), pp. 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), pp. 249-263.

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), pp. 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), pp. 523-553.

Rosenthal, R. & Jacobson, L. (1968). Pygmalion in the Classroom: Teacher Expectation and Pupils' Intellectual Development. New York: Holt, Rinehart and Winston, Inc.

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), pp. 429-444.

Slavin, R. E. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. *Review of Educational Research*, 57, (3), pp. 293-336.

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), pp. 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, pp. 57-67.