Marion M. de Lemos reports on a longitudinal study which asked: Does multi-age grouping make a difference to the students in terms of their learning environment or their social adjustment?

INTRODUCTION

In recent years a policy of multi-age grouping at the lower primary level has been adopted by a number of schools and school systems, both in Australia and overseas. This trend is based on the belief that multi-age grouping encourages a more developmentally appropriate approach to the teaching of young children, leads to greater recognition of the individual differences between children in terms of their level of development and their language and conceptual skills on entry to school, and provides a more natural learning environment that optimises children’s social and academic learning through interaction with children of different ages and at different levels of development.

Multi-age grouping is not a new concept. It has always existed in small country schools where enrolments are insufficient to establish separate classes for each age group. There is, however, a distinction between composite classes formed for purposes of administrative convenience and multi-age classes which are formed to implement a particular approach to teaching.

The concept of the non-graded school as an alternative to the traditional age-graded school structure became popular in the USA in the 1960s and 1970s. However, this form of grouping was implemented in only a small proportion of schools in the USA, and over time the non-graded class became associated with other innovations in the classroom organisation and teaching practice, such as individualised instruction, team teaching and open classrooms.

Mixed age or vertical grouping was also found in many British primary schools in the 1960s and 1970s, with this form of grouping being more common at the infant level than at the junior level of primary schools. There was, however, a decline in vertical grouping over the period 1973 to 1980, following the publication of a survey by HM Inspectors which indicated that the achievement level of students in mixed-age classes tended to be lower than that of their peers in single-age classes.

The impetus for the adoption of multi-age grouping in recent years has focused on younger children at the pre-school or early primary level, and is associated with the principles of developmentally appropriate practice. However, in some countries, an increase in the number of multi-age or mixed-level classes is associated with falling enrolments and economic and political factors rather than a consideration of children’s needs, and is seen as a problem to be overcome rather than an advantage to students.

The most recent review of the research evidence on the effects of multi-grade and multi-age classes is the best-evidence synthesis published by Simon Veenman in 1995. In this review, Veenman examined the evidence relating to both cognitive and non-cognitive effects of multi-grade and multi-age classes.

In terms of cognitive effects, Veenman found that comparisons between multi-grade and single-grade classes, as well as between multi-age and single-age classes, indicated no consistent differences between these two types of classroom organisation in terms of achievement (median effect size = 0.00 and −0.03 respectively). In terms of non-cognitive effects, Veenman found that comparisons between multi-grade and single-grade classes and between multi-age and single-age classes indicated no consistent differences between these two types of classroom organisation on various measures of attitudes to school, self-concept, motivation and social adjustment, although effect sizes were on average positive, indicating some tendency for more positive non-cognitive outcomes in the case of children in multi-grade or multi-age classes (median effect size = 0.10 and +0.15 respectively).

On the basis of this evidence, Veenman concluded that there is no empirical evidence that student learning may suffer in multi-grade and multi-age classes. Veenman’s conclusions have, however, been challenged by Mason and Burns, who argue that Veenman has ignored two important factors that make his ‘no difference’ conclusion problematic. These factors are selection bias and lower-quality instruction.
In terms of selection bias, Mason and Burns argue that evidence from the literature indicates a tendency for more able and better adjusted students, and for better or more experienced teachers, to be allocated to multi-grade and multi-age classes. In terms of quality of instruction, Mason and Burns argue that evidence from the literature indicates that teaching in multi-grade and multi-age classes is associated with more time required for preparation, less instruction time (per grade level group), less time for attention to individual students, greater demands for class management and greater teacher stress – all of which lead to negative outcomes in terms of curriculum coverage and adaptive assistance.

Mason and Burns argue that the finding of ‘no difference’ between multi-grade or multi-age classes and single-grade or single-age classes is, in effect, due to the compensatory effect between these opposing trends (that is, better students and teachers together with lower-quality instruction). On the basis of their own review of the evidence, they conclude that multi-age grouping has at least a small negative effect on achievement, as well as potentially negative effects on teacher motivation.

Regardless of whether one accepts Veenman’s or Mason and Burns’s interpretation of the data, it would seem clear that the research evidence indicates no clear difference between students’ achievements and social adjustment according to whether they are placed in a multi-age or a single-grade class.10

In Australia, support for the adoption of mixed-age grouping at the lower primary level was expressed in a number of government reports and discussion papers in the late 1980s and early 1990s, with this form of grouping being seen as having positive outcomes both in terms of the children’s social development and of learning outcomes.11 While the more recent research has challenged these assumptions, it was in the context of the earlier, more favourable view of mixed-aged grouping that policies favouring multi-age grouping were adopted in a number of Australian school systems.

EVALUATION OF THE VICTORIAN MULTI-AGE PILOT PROJECT

In Victoria, a pilot project on multi-age grouping was initiated in 1994, following the 1992 Report of the Ministerial Review of School Entry Age in Victoria, which recommended the adoption of more flexible grouping practices in the early years of schooling.12 This project provided an opportunity to monitor the impact of this form of grouping on children’s social development and learning outcomes.

An independent evaluation of the Victorian multi-age project was undertaken by researchers at the Australian Council for Educational Research, the University of Melbourne, and Monash University.

The design of the evaluation study involved the collection of data at the commencement of the study, to determine levels of achievement prior to the introduction of multi-age grouping; during the course of the study, to monitor the progress of children in multi-age and single-grade classes as they progress through the early years of school; and at the conclusion of the study, to determine whether multi-age grouping had any impact on student achievement over the three years of the pilot project.

Baseline measures were administered in the first year of the study to samples of children drawn from both the Prep grade (the first year of school) and Year 3. The children in the Prep sample were then reassessed each year on standardised measures of reading and mathematics, as well as on teacher checklists of progress in reading and mathematics. Information on the children’s social skills and academic self-esteem were also obtained in the third and fourth years of the study; and in the fourth and final year of the study (1997), when the children were in Year 3, they were assessed on the same measures of literacy and numeracy that had been administered to the 1994 Year 3 sample.14

Overall, the results of this study indicated no significant differences between the scores of the pilot and control samples on the baseline measures administered at the start of the study, or on the measures of achievement in reading and mathematics that were administered during the course of the study. There was, however, some tendency for the children in the control sample to score higher on the measures of reading, and for this tendency to increase over time.

In an educational context, it is the size of the difference rather than its statistical significance that is usually of more consequence. For this reason, effect sizes are increasingly being used as a way of presenting results in terms that are educationally meaningful. Effect sizes express differences between groups in terms of standard deviation units. An effect size of 0.20 of a standard deviation is generally interpreted as a small effect, an effect size of 0.50 is generally interpreted as a moderate effect and an effect size of 0.80 is interpreted as a large effect.15

In order to examine the pattern of responses on the various measures administered, and also to compare the findings of this study with the findings reported in the literature, effect sizes for the differences in score between the pilot and control samples were calculated by subtracting the mean raw score of the control group from the mean raw score of the pilot group, and dividing this difference by the standard deviation of the control group. Positive effect sizes indicate higher scores for the pilot sample, while negative effect sizes indicate higher scores for the control group.

Effect sizes for the various measures of academic achievement included in the study are shown in Table 1. From this table it can be seen that the differences between the two samples in the middle of the first year of school was negligible (–0.03 on
the AGS Language Profile and 0.09 on the Cognitive Profile). On the standardised measure of reading achievement, however, there is a clear trend for the differences between the two samples to increase from Prep to Year 3, with the negative values of the effect sizes indicating that the mean scores of the pilot sample are consistently lower than the mean scores of the control sample. In the case of the teacher assessments, a similar trend is found for the students in the pilot sample to receive lower assessments for progress in reading than in the case of students in the control sample.

**Table 1. Summary of effect sizes for differences between pilot and control samples on measures of academic achievement, by type of measure and year of schooling, for 1994 Prep sample**

<table>
<thead>
<tr>
<th>Measure</th>
<th>1994 Prep sample</th>
<th></th>
<th></th>
<th></th>
<th>Fourth year (Year 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Middle of first year (Prep)</td>
<td>End of first year (Prep)</td>
<td>Second year (Year 1)</td>
<td>Third year (Year 2)</td>
<td></td>
</tr>
<tr>
<td>AGS ESP:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language profile</td>
<td>-0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitive profile</td>
<td>0.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standardised test</td>
<td>-</td>
<td>-0.10</td>
<td>-0.21</td>
<td>-0.21</td>
<td>-0.24</td>
</tr>
<tr>
<td>Teacher assessment</td>
<td>-0.18</td>
<td>-0.21</td>
<td>-0.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mathematics:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standardised test</td>
<td>-</td>
<td>-0.05</td>
<td>0.05</td>
<td>0.08</td>
<td>-0.11</td>
</tr>
<tr>
<td>Teacher assessment</td>
<td>0.15</td>
<td>0.32</td>
<td>0.08</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In the case of mathematics, the results are less clear, and there are inconsistencies between the teacher assessments and the standardised measures. While the teacher assessments tend to favour the students in the pilot sample, the standardised measures indicate little if any difference between the two groups, although at the Year 3 level there is a negative effect size indicating that the control sample is scoring at a marginally higher level than the pilot sample.

In comparing the raw scores of the 1994 Year 3 sample with those of the 1997 Year 3 sample on the ACER Tests of Basic Skills, it was noted that there had been some improvement from 1994 to 1997. In the case of the literacy test, there was an increase in mean raw score of just over two points, while in the numeracy test, there was an increase of about one point. This increase was, however, found in both the pilot sample and the control sample, so this overall improvement in score cannot be attributed to the effect of class organisation. Given the fact that the 1994 Prep sample had been assessed annually on standardised measures of reading and mathematics over the three years of the project, it is possible that this increase in score is a result of greater familiarity on the part of the children in responding to this type of assessment.

These results therefore indicate little if any effect of multi-age grouping on children’s academic achievement, and are consistent with the research evidence that indicates, if anything, a small negative effect of multi-age grouping on cognitive outcomes.

**Multi-age grouping**

---

*Spring 2001 • Issue 25*
NON-COGNITIVE EFFECTS

Effect sizes for the various measures of social development included in the study are shown in Tables 2 and 3. Table 2 summarises the results on the three measures of social development provided by the AGS Early Screening Profiles, which were collected at the start of the study (a measure of social and self-help skills, completed by the parents of the children in the first year of the study; a measure of speech articulation, administered in conjunction with the Language/Cognitive Profile, and a measure of behaviour as observed during the administration of the Cognitive/Language Profile), as well as the measure of academic self-esteem, administered to the 1994 Prep sample in the third year of the study (the 1996 Year 2 group). Since there was a gap in the data on social skills from Prep to Year 2, additional samples of children in Prep and Year 1 were also assessed in the third year of the study on the measure academic self-esteem. These additional data were designed to provide cross-sectional data on the development of academic self-esteem over the first three years of schooling in single-grade and multi-age classes.

Table 2. Summary of effect sizes for differences between pilot and control samples on the Behavioural Academic Self-esteem Rating Scale: for 1996 additional cross-sectional samples of children assessed in June 1996

<table>
<thead>
<tr>
<th>Measure</th>
<th>1994 Prep sample</th>
<th>1996 1st year (Prep)</th>
<th>1996 2nd year (Year 1)</th>
<th>1996 3rd year (Year 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGS ESP:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-help/social profile</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Articulation survey</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behaviour survey</td>
<td>0.18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behavioural academic self-esteem:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total BASE score</td>
<td>-0.24**</td>
<td>0.06</td>
<td>0.17*</td>
<td></td>
</tr>
<tr>
<td>Subscale scores:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student initiative</td>
<td>-0.35**</td>
<td>0.03</td>
<td>0.14*</td>
<td></td>
</tr>
<tr>
<td>Social attention</td>
<td>-0.09</td>
<td>0.05</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>Success/failure</td>
<td>-0.07</td>
<td>0.07</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>Social attraction</td>
<td>-0.15</td>
<td>0.11</td>
<td>0.16*</td>
<td></td>
</tr>
<tr>
<td>Self-confidence</td>
<td>-0.18*</td>
<td>0.05</td>
<td>0.08</td>
<td></td>
</tr>
</tbody>
</table>

* Indicates that differences in mean scores are significant at the 0.05 level.
** Indicates that differences in mean scores are significant at the 0.01 level.

Results on the measure of social skills administered in the third and fourth years of the study are shown in Table 3. The rating of social skills was completed by the teachers of the 1994 Prep sample in their third and fourth years at school (Years 2 and 3), with additional information based on parent and student ratings collected in the fourth year of the study (when the 1994 Prep sample was in Year 3).

As seen in Table 2, results on the measures of social development and behaviour collected in the first year of the study indicated no significant differences between the pilot and control samples in their self-help and social skills, speech maturity or behaviour patterns at the start of the study. If anything, the tendency was for the children in the pilot sample to be rated more positively by their teachers in terms of their social behaviour. When this sample was reassessed in 1996 on the measure of academic self-esteem, children in the pilot sample scored at a significantly higher level than children in the control sample on the overall measure of academic self-esteem, as well as on two of the subscales of this measure (student initiative and social attraction). These data on their own would appear to indicate a positive effect on social development for the children in multi-age classes. However, the cross-sectional data on children at the Prep and Year 1 level on the measure of academic self-esteem, collected at the same time as the longitudinal data on the 1994 Prep sample, suggest a rather different interpretation of these results. When the results across the three year levels are considered, the trend is for children in multi-age classes to score at a lower level than children in single-grade classes at the Prep level, when the children in multi-age classes are younger than the majority of the children in their class group, and at a higher level than children in single-grade classes at the Year 2 level, when the children in multi-age classes are older than the majority of the children in their class group. An effect of relative age in grade was also noted within each grade level, with the younger children within a grade level tending to score at a lower level than older children within a grade level. These results suggest that the pattern of scores from Prep to Year 2 of children in multi-age versus single-grade classes on this measure may be attributable to an effect of relative age in grade rather than to an effect of multi-age grouping as such.
Table 3. Summary of effect sizes for differences between pilot and control samples on the teacher, parent and student forms of the Social Skills Rating System, 1996 and 1997

<table>
<thead>
<tr>
<th>Social skills rating domains and subscales</th>
<th>1996 Teacher form</th>
<th>1997 Teacher form</th>
<th>1997 Parent form</th>
<th>1997 Student form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social skills</td>
<td>0.02</td>
<td>–0.30**</td>
<td>0.03</td>
<td>–0.14</td>
</tr>
<tr>
<td>Problem behaviours</td>
<td>0.11</td>
<td>0.15</td>
<td>0.07</td>
<td>–</td>
</tr>
<tr>
<td>Academic competence</td>
<td>0.21**</td>
<td>–0.13</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Social skills subscores:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cooperation</td>
<td>0.05</td>
<td>–0.31**</td>
<td>0.06</td>
<td>–0.10</td>
</tr>
<tr>
<td>Assertion</td>
<td>0.17</td>
<td>–0.18</td>
<td>–0.14</td>
<td>–0.15</td>
</tr>
<tr>
<td>Responsibility</td>
<td>–</td>
<td>–</td>
<td>0.06</td>
<td>–</td>
</tr>
<tr>
<td>Empathy</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–0.14</td>
</tr>
<tr>
<td>Self-control</td>
<td>0.00</td>
<td>–0.32**</td>
<td>0.06</td>
<td>–0.12</td>
</tr>
<tr>
<td>Problem behaviours subscores:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Externalising problems</td>
<td>0.15</td>
<td>0.08</td>
<td>0.03</td>
<td>–</td>
</tr>
<tr>
<td>Internalising problems</td>
<td>0.04</td>
<td>0.08</td>
<td>0.15</td>
<td>–</td>
</tr>
<tr>
<td>Hyperactivity</td>
<td>0.05</td>
<td>0.21</td>
<td>0.04</td>
<td>–</td>
</tr>
</tbody>
</table>

** Indicates that differences in raw score are statistically significant at the 0.01 level.

In the case of the measure of social skills administered to the 1994 Prep sample in Years 2 and 3, there were some differences between the pilot and the control samples, but these differences were not consistent over time. However, the direction of the differences was consistent with the hypothesised effect of relative age in grade. That is, the children in the multi-age classes tended to score higher than the children in the control sample on the measures of social skills at the Year 2 level, when they were significantly older than the majority of the children in their class group, and lower than the children in the control sample at the Year 3 level, when the majority of the children in the pilot sample were in composite Year 3/4 classes, and therefore relatively younger than the majority of the children in their class group. These results, again, point to relative age in grade as a more likely factor in explaining the different patterns of results on the measures of social skills administered to the pilot and control samples in Years 2 and 3.

Consistency with other research findings

These results are therefore consistent with the overseas research evidence, indicating that multi-age grouping has no significant impact on either academic achievement or social development. A weakness in most studies of the effects of multi-age or multi-grade classes has been the lack of data on the children’s levels of achievement prior to entry to the multi-age or multi-grade class. The longitudinal nature of this study and the fact that the comparability of the samples of children in the pilot and control groups was established at the start of the study is therefore an important strength of the study.

In the case of the study described in this article, the issue of selection bias has also been controlled, in that there is no allocation of individual students to multi-age or single-grade classes on the basis of personal characteristics. All children at the Prep to Year 2 level at the pilot schools who participated in this project were allocated to multi-age classes, while all children in the control schools were allocated to non-multi-age classes. These non-multi-age classes were in most cases single-grade classes, but did include some composite classes which were combined for administrative purposes, and generally maintained a distinction between the different grade level groups within the same classroom.

Data from the teacher questionnaires administered as a part of the evaluation study support the findings from the literature that teaching in multi-grade and multi-age classes is associated with more time required for preparation and greater demands on teachers’ organisational and class management skills. However, the teachers in the pilot schools were provided with a greater degree of support and professional development than was the case for teachers in the control schools, and it is possible that this factor compensated to some extent for the negative effects associated with the greater difficulties experienced by teachers in managing a multi-age classroom, thus introducing a compensatory effect as suggested by Mason and Burns.

The finding that students in the control sample tended to score higher on the standardised measures of reading administered in their second, third and fourth years of schooling, as well as on the teacher assessments of progress in reading administered at the end of the first, second and third years of schooling (effect sizes from –0.10 to –0.24), is consistent with Mason and
VERTICAL GROUPING IN THE UK: IS IT AN ISSUE?

How widespread is vertical or multi-age grouping in the UK? Is it adopted for administrative reasons or educational reasons? Is it found only in schools where numbers are too small for single-grade classes, or where enrolments across year levels are uneven, or is it adopted by schools because it is seen as providing a benefit for pupils? Is it linked to school entry policies (termly entry as opposed to single entry at the beginning of the school year)? Are vertically grouped classes more common at the infant level (Reception to Year 2), or are they found throughout the primary level? To what extent is vertical grouping linked with other forms of grouping within the class (ability grouping, setting, streaming, open classrooms)? Is vertical grouping linked with specific types of teaching approach (more structured versus less structured, whole language-based versus skills-based teaching, implicit versus explicit teaching of phonics)? What, if any, effect does vertical grouping have on student outcomes?

A search of the British literature yields little in reply to these questions. In his comprehensive review of the cognitive and non-cognitive effects of multi-grade and multi-age classes, published in 1995, Veenman lists only three British studies among the 56 studies identified as providing information on the effects of multi-grade or multi-age grouping. These three studies include the 1978 survey of primary education in England, undertaken by HM Inspectors of Schools, the 1980 Galton and Simon study on progress and performance in the primary classroom, and a 1966 unpublished Master’s thesis on a comparison of vertical grouping and horizontal grouping in the infant school, undertaken at the University of Manchester by Mycock.¹

As Veenman points out in his review, there are a significant number of multi-grade classes in most school systems, both among Western industrialised nations and in developing countries. These classes are found in both urban and rural areas, as a consequence of low enrolments in thinly populated areas or shifting enrolments in urban and suburban areas. Veenman also points to the effects of financial cuts leading to the reduction of teacher numbers and falling enrolments due to declining birth rates as factors leading to an increase in the number of multi-grade classes.

According to a 1982 report by HM Inspectorate, almost half of the new teachers in England and Wales had their first posts in multi-grade classes. And on the basis of a survey of 642 schools in England and Wales carried out in 1984, Walsh et al. reported that 40 per cent of teachers said they had experienced an increase in multi-grade teaching as a result of falling enrolments, and a further 15 per cent said declining enrolments might lead to an increase in multi-grade teaching in the future.

More recent data on the prevalence of multi-grade classes in the UK does not seem to be available. However, it seems likely that this form of class grouping is commonly adopted not only in small rural schools, but also in urban and suburban areas, either for administrative reasons or on the basis of the assumed benefits of this form of grouping.

Veenman’s review of the research evidence indicated no consistent negative effects associated with multi-grade or multi-age grouping. However, this interpretation of the evidence has been challenged by Mason and Burns, who argue that Veenman’s finding of no effect fails to address the issue of the possible compensatory effects of selection bias and lower-quality instruction in multi-grade and multi-age classes.

The evidence from Australia, reported in this study, indicates a consistent negative effect for students in multi-age classes on measures of reading, but not on measures of mathematics. Effects on social development and self-esteem are also noted, but in this case the effects seem to vary according to whether the children in the comparison multi-age class are in the older or the younger group within the multi-age class, thus suggesting a relative age effect rather than an effect of multi-age grouping as such.

This evidence suggests the need for more attention to the possible effects of vertical grouping in the UK, including the collection of more systematic data on the prevalence of this form of grouping, the reasons for its adoption, its impact on teaching practices and its possible impact on student outcomes.

¹ All references to the British research in this area are quoted from Veenman (1995). Details of this and other relevant references are provided in the article.
Burns’s tentative conclusion, based on their analysis of the relevant evidence, that multi-age grouping has a small negative effect on achievement (of the order of about –0.10). This effect was not, however, found in the case of mathematics, where the differences between the two groups varied both over time (from the end of the Prep year to Year 3) and according to the nature of the assessment (teacher assessments versus standardised tests).

In the case of social development, the results of this study point to relative age in grade as a more important factor affecting social behaviour than classroom organisation. That is to say, there was a tendency for the younger children within a class group to score at a lower level than older children, regardless of whether the class was a single-grade or a multi-age group. In view of the fact that multi-age grouping is intended to provide a more advantageous social climate for the school beginner, this finding raises questions about the assumption that multi-age grouping enhances the social development and self-confidence of young children in their first year of schooling.

IMPLICATIONS FOR SCHOOLS

The findings from this study can be interpreted in both a positive and a negative way. In a negative sense, it could be argued that the lack of any clear evidence for any advantage of multi-age grouping, either in terms of student learning or in terms of social adjustment, raises questions about the rationale for introducing this form of grouping in preference to single-grade grouping, particularly in view of the greater demands on teacher time and skills that this form of grouping requires.

In a positive sense, these results indicate that in cases where some form of multi-age grouping is required for administrative reasons, this form of grouping will not necessarily disadvantage children.

More importantly, the results of the overseas research as well as the Victorian study point to the critical role played by the teacher in any form of class grouping or class organisation. Ultimately, it is the quality of the teaching programme and the effectiveness of the teacher that determines the programme’s outcomes.

Notes
Marion M. de Lemos is a Senior Research Fellow at the Australian Council for Educational Research, E-mail: delemos@acer.edu.au

1 For example, see:

2 For the ‘open schools’ of the 1970s, see:

3 For British examples, see:

4 Decline in vertical grouping:

5 Comparing achievement of mixed-age and single-age classes:

6 For example, see:
KATZ, L. G., EVANGELOU, D. and HARTMAN, J. A. *The Case for Mixed Age Grouping in Early Education*. Washington, DC: National Association for the Education of Young Children and
and

7 See:


Recent review of effects of multi-age classes:

Challenge to Veenman:

It is of interest to note that data from the New Zealand study of competent children at age eight are consistent with these findings. While initial analyses indicated a small but significant tendency for children in single-grade classes to score at a higher level than children in composite classes, further analyses indicated complex interactions between various other factors affecting children’s competencies at age eight; see:

See, for example:


Recommend flexible grouping:

Further information on this study:

Tests administered in the study:

Interpretations of statistical significance:

Demands on class management skills:

Acknowledgements
This study was undertaken in conjunction with Dr Christine Ure, of the University of Melbourne, and Dr Elizabeth Mellor, of Monash University, with the support of the Victorian Department of Education and the collaboration of the staff and students at the schools participating in the study. Their contribution to the study is gratefully acknowledged.

Copying Permitted
The major part of this article is reprinted with permission from set Number 2, 1999, published by the New Zealand Council for Educational Research and the Australian Council for Educational Research. Copyright is held by NZCER and ACER, who grant to all people actively engaged in education the right to copy it in the interests of better dissemination. Please acknowledge the source.