This is the published version:

Doig, Brian 2005, Analyzing Australian middle years students' preferred mathematics classroom practices, in ICMI-EARCOME 3: The third East Asia Conference on Mathematics Education proceedings, East China Normal University, [Shanghai, China], pp. 1-7.

Available from Deakin Research Online:

http://hdl.handle.net/10536/DRO/DU:30014519

Reproduced with the kind permission of the copyright owner.

Copyright: 2005, The Author
Analyzing Australian Middle Years Students’ Preferred Mathematics Classroom Practices

Brian Doig
Deakin University
Australia

Abstract

The Middle Years of schooling have become an issue for mathematics teachers and educators, and calls for the reform of this period of schooling are frequent. However, the suggested reforms appear to be divided in their views of what would be best for these students. This paper provides a brief overview of what some Middle Years students say about their needs. Some 1500 students in the Middle Years of schooling (Years 5 to 9) were surveyed about their preferred mathematics classroom activities: the methods and results of the survey may be extremely relevant to those contemplating Middle Years reform.

Keywords: curriculum reform, assessment, students’ views, classroom practices

Introduction

The Middle Years of schooling, generally defined as Years 5 to 8, has become a focus of attention for Australian educators, with concern for the achievement of these students raised by the recent results of international studies. Luke et al. (2003) and Lokan and her colleagues (2001), report that academic performance in mathematics drops in the transition year from primary to secondary school, and this regardless of the student’s age when this transition occurs. This ‘dip’ in achievement is not confined to mathematics, nor just to Australian students (Felner, Shim, Favazza, & Seitsinger, 2000), although the issue is not universal (Stigler & Hiebert, 1999). This raises questions about the differences in pedagogical practices between teaching cultures, and also suggests that the current situation, in Australia and elsewhere, is neither inevitable nor insoluble.

The Cockcroft Report, from the UK (Cockcroft, 1982), suggested that by age fourteen there is up to a 7-year difference in students’ achievement in mathematics, while Hill and his colleagues (Hill, Holmes-Smith, & Rowe, 1993; Sullivan, 2003) noted that the lowest attaining students do not appear to progress academically beyond Year 4. In the reports on the achievements of Australian Middle Years students in international comparative studies, Lokan and her colleagues (Lokan, Ford, & Greenwood, 1996, 1997; Lokan et al., 2001) have documented the apparent ‘gaps’ in achievement of Australian students, and the difficulty Australian students have with more complex mathematical procedures and more abstract mathematical ideas (see, Doig, 2001 for a summary of these findings). Research also shows that there is a growing gap between those students who see themselves as successful at school and those who do not (Lokan et al., 2001).

Ways of overcoming these problems need to be explored. Cogan and Schmidt (2003) recommended that the mathematics curriculum should be coherent across the grade levels, for all topics, and that fewer topics should be taught in the Middle School, but not all researchers agree with Cogan and Schmidt’s point of view (see Felner et al., 2000, for example). Several researchers have remarked on the level of mathematical learning demanded of Middle Years students. That is, the level and quality of the intellectual demands on our students (see, for example, Luke et al., 2003).

Australian Middle Years students’ preferred mathematics classroom practices
Norton and Lewis (2000) produced a special report on Middle Grades reform in the United States for Phi Delta Kappa, in which they describe two successful school reform strategies. In the first successful reform strategy, it was reported that team planning (Middle Years teachers planning together); authentic teaching and learning (curriculum co-ordination, co-ordination of assessments, contact with parents and resource staff); and parent involvement (student achievement, homework social problems) were major factors in producing student improvement in reading and mathematics. Typically, the more co-ordination across grade levels, the higher the academic gains. The second example of a reform endeavour, was sited in a single school, where teachers began a series of ‘self-study’ professional development experiences, which spurred a continuing programme of professional development, and the development of strategies to improve students’ academic achievement. Other features of the school were the use of multi-age classes, and teachers staying with the same students as they move up grade levels. The success of this multi-faceted approach was evident in the increased achievements of the students in this school.

On a larger scale, Felner and his colleagues (2000) reported findings from a project on high performing learning communities. This research group employs a set of assessment instruments completed by teachers, students, and parents, known as the High Performance Learning Community (HiPlaces) Assessment. This suite of assessments gauges the degree to which a school implements (full-, partial-, or low-implementation) the nine dimensions of High Performance Learning Communities. These nine dimensions include a high quality core curriculum that is rigorous, combined with high expectations of all students, and the development of small, personalized, learning communities, among others. Students in the 31 project schools were also part of state-wide assessment programs, and their results were that high implementation schools had a mean score of 298, partial implementation schools a mean of 279, and the low implementation schools a mean of 248, while the state average was 250.

In the Australian state of Victoria, the Middle Years Research and Development project (MYPRAD) used ‘component mapping’ to audit teachers’ current practices. These components provide a picture of what Middle Years practice looks like, and guidelines for teachers in terms of alternative teaching strategies. The component mapping exercise involves individual teachers working through the components, with a significant other, to establish their level of practice with respect to each of the five components and their sub-components (25 elements altogether) (Department of Education and Training Victoria, 2004). Tytler (2004) reported the analysis of the data from the teachers in the project who were component mapped. From this mapping, the factors that teachers saw as needing more attention were listening to students, a more flexible curriculum, better links with the community, more variety in teaching strategies, and a move away from teacher-centred teaching.

However, what aspects of classroom practice are seen by the students themselves, as progressing their learning of mathematics, remains an un-answered question.

The survey

As part of the Australian Research Council funded project, Improving Middle Years Mathematics and Science (IMYMS), nearly 1500 students in Year 5 through to Year 9 were surveyed (Primary N = 713; Secondary N = 744). Surveys lend themselves to a range of research applications (Anastasi, 1990), although a common use in educational research is for collecting attitude or opinion data. In this case, the survey was a rating scale, with stimulus statements that require the students to indicate, or rate, their position with respect to the ‘attitude object’ specified in statements about the helpfulness of various classroom practices.

The method devised by Likert (1932) was used in the IMYMS student survey. For example, a student may be asked to indicate their view of the efficacy of a classroom practice such as “Taking part in class discussion”. The student may indicate their response by a tick for the response statement such as ‘Helpful’. Table 1 shows part of a typical Likert-type survey.
Table 1: Responses to a Likert-type survey form.

<table>
<thead>
<tr>
<th>Item</th>
<th>Statement</th>
<th>Very Helpful</th>
<th>Helpful</th>
<th>Not sure</th>
<th>Unhelpful</th>
<th>Not at all Helpful</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Taking part in class discussions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Going on excursions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Copying notes from the board</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The IMYMS survey had two aspects. The first focused on features of the classroom environments in which the students found themselves: the survey focused on features such as the degree to which teachers encouraged mathematical dialogue between students, or used investigations as learning experiences. The second focus of the survey was students’ perspectives on, or attitudes to, teachers’ particular classroom practices: for example, the usefulness of excursions for their learning of mathematics. The complete set of questions for this second focus is shown in Figure 2 below.

Surveys such the IMYMS survey are generally known as rating scales, and while it is true that there is an apparent simplicity in rating scales, there are, as with all data collection and analysis approaches, some fundamental considerations that need addressing.

The underlying concept behind a rating scale is that it is a measure of some latent, or hidden, trait or characteristic, or in this case, attitude or viewpoint. This means that, in a sense, the rating scale performs a duty in a manner not unlike an achievement test. However, it is an achievement test where questions may have a series of correct answers (or where there is a continuum of increasing sophistication in responses to an item: see Tapping Students Science Beliefs for an example of this type of achievement assessment (Doig, 1995; Doig & Adams, 1993; Doig & Adams, 1991). In this sense, the IMYMS survey of student attitudes on the usefulness of various classroom practices is a rating scale where ‘usefulness’ is the underlying trait.

**Analysis**

The responses to the twenty-four rating scale statements focused on classroom practice were analysed using a Master’s Partial Credit Model (Bond & Fox, 2001; Masters, 1982) by the Item Response Theory (IRT) software Quest (Adams & Khoo, 1993). The Item Response Theory approach does not assume, like traditional analyses, that it is equally easy, or hard, to agree to every stimulus statement. This approach assumes that a respondent may have differing attitudes to, or views about, different aspects of the attitude object, as defined by the stimulus statement. It may be much easier to agree that “Going on excursions” is “Very helpful” than to agree that, “Doing homework” is “Helpful”, for example.

The IRT and traditional analyses differ also in their view of the difference between response categories. The traditional approach assumes that it requires the same amount of change in viewpoint to move to from one response category to another. That is to say, that the difference in change to one’s attitude or viewpoint, required to move from “A bit helpful” to “Helpful”, for example, is the same as that required to move from “Not helpful” to “A bit helpful”. On the other hand, the IRT approach assumes that differences in amount of change to one’s perspective may exist, and allows the boundaries between categories to be specified by the set of response data. The results of this form of analysis is an equal interval scale that provides clear indications of the ease, or difficulty, that respondents have in agreeing with the stimulus statements of the survey. Further, the interval scale constructed has a unique property: both the difficulty of selecting a particular category of response, and the
“attitude” of each respondent, are measured on the same scale. This enables us to estimate the likelihood with which a respondent would select a particular response category for any item once we know their total score.

Results

As can be seen from Figure 2, there are three response categories (Very helpful, A bit helpful, and, Not helpful), which means that there are only two thresholds, or boundaries on the ‘helpfulness’ scale. The values provided by the Quest IRT analysis are stated in ‘logit’ units, that range from about –4.0 to about 4.0 in this case. Thus our “helpfulness” scale has a range of about 8 logits.

The Quest analysis also provides an estimate of the overall attitude or viewpoint that each respondent has exhibited through their responses. As mentioned above, these estimates are on the same scale as the category thresholds. This is a very important point when it comes to reporting the findings of the analysis. For the survey that we are using here, the mean for the Primary school students is a raw score of 33 (0.96 logits), and the mean for the Secondary school students is a raw score of 25 (0.11 logits). The position of these means is marked in Figure 2 by two labelled, vertical lines.

The ‘bar’ beside each statement shows the three categories of response: the shaded section represents the “A bit helpful” category, while the two end sections represent, left to right, “Very helpful” and “Not helpful”. The length of the ‘bar’ sections indicates the category of response likely for a specific range of student raw score totals. In other words, knowing a student’s total raw score is sufficient to know their likely response to every item. This is shown visually by simply drawing a vertical line upwards, from a student’s total raw score position from Question 24 to Question 1. Usually, a raw score scale is provided at the bottom, or top, of the set of questions; in Figure 2, for simplicity, the total scores scale has been omitted. The raw score scale is based on 24 items each with a maximum of 3 score points — a possible total score of 72.

Interpretation

The results of the IRT data analysis are shown in the report in Figure 2. In this form of report, a visual approach is taken, as this provides an immediacy unavailable in a table of figures.

The vertical mean score line for the Secondary students indicates their most likely response category for each item. In Item 1, “Taking part in class discussion” for example, the most likely response of a Secondary student is “A bit helpful” while for Item 5, “Working in small groups”, it would be “Very helpful”. For Primary students, the most likely response for Item 12, “Going on excursions”, is “Not helpful” while for Item 18, “Giving a talk to the class”, would be “A bit helpful”.

In some cases, the mean score line passes through the line separating categories, or is very close to this position, indicating that caution must be exercised in interpreting which category is the most likely (see, for example, Item 19 for Primary students).
<table>
<thead>
<tr>
<th></th>
<th>Very helpful</th>
<th>A bit helpful</th>
<th>Not helpful</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Taking part in class discussions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Doing exercises or answering questions from a book</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Doing hands-on activities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Doing homework</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Working in small groups</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Doing worksheets</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Listening to the teacher explain ideas</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Playing games</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Doing investigations of my own choice</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Using computers or calculators</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Watching the teacher show us how to do things</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Going on excursions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Copying notes from the board</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Talking in class about things in the news</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Asking questions about things that interest me</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Searching and collecting information using the Internet or CD-ROM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Doing projects in my local community</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Giving a talk to the class</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Listening to a visiting speaker</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Being able to choose how I present things</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Watching videos</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Writing my thoughts about what I've learnt in a diary or journal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Doing activities that challenge me to think</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Searching for information using library books</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 2: The IMYMS rating scale.**
Differences between Secondary and Primary students’ mean responses are immediately visible on this format of report. For example, for Item 5, “Working in small groups”, the secondary students are most likely to indicate that this is very helpful to their learning of mathematics, but the Primary students’ responses indicate that they view working in small groups as only a bit helpful. A similar situation is revealed for other items (see, for example, Item 8, “Playing games”, Item 17, “Doing projects in my local community”, and Item 22, “Writing my thoughts about what I’ve learned in a diary or journal”).

In this paper, mean scores for groups of students have been used for illustrative purposes, but similar interpretations can be made for other groups, such as males and females, and even for individual students using their total score on the twenty-four items. The principle remains the same.

The difference between Primary and Secondary mean scores shows that the Primary students were less favourable in their views that the suggested teaching practices were as helpful to them as to their Secondary counterparts. In most instances, they were about a category lower in their opinions than the Secondary students, although Item 2, Item 10, and Item 20 were exceptions to this.

Finally the length of the sections representing the extent of the opinion categories would suggest that some classroom practices are seen as being very helpful by a greater proportion of students (a notable example is Item 5). On the other hand, Item 1 would appear to be rejected as being very helpful by the majority of students. The least helpful classroom practices would appear to be Item 3, “Doing hands-on activities”, Item 9, “Doing investigations of my own choice”, and Items 16 and 24, both of which are similar in nature.

**Conclusions**

It is unarguable that it is important to take students’ preferences, or views, into account when planning teaching and learning experiences for them. However, there appears to be little evidence of the research into Middle Years reform taking this into account. The **Improving Middle Years Science and Mathematics** project student survey is an attempt to make visible these student preferences, and report them in a way that is accessible to the teachers in the project.

As presented here, an examination of the preferences of this group of students reveals that some classroom practices are believed by the students not to be helpful to them, and it is this information that teachers can make use of in planning learning experiences for their students. The differences between the preferences of Primary and Secondary students (and possibly other sub-groups too) would suggest that so-called generic teaching approaches may not be effective for all students, and produce differentiated learning gains.

Although surveys of student viewpoints are useful, the format in which the results are presented is critical in making the results immediate and comprehensible to teachers and students. The use of Item Response Theory analysis clearly provides reports that are clear and meaningful, and the IMYMS report, shown here, is one format in which this can be achieved.
References


Likert, R. (1932). A technique for the measurement of attitudes. Archives of psychology (Number 140).


Literacy and Numeracy development of target group students in the middle years of schooling. Nathan, Queensland: DEST Clearinghouse, Griffith University.


