## Chapter 3

# RESEARCH DESIGN

Several research questions emerged from the review of the literature in the first two chapters. These were formulated into four research themes which were aimed at providing answers to the general question of the applicability of the Mayberry test.

The objective of this chapter is to describe the development of the written test, and the methodological considerations incorporated into the study. To assist in this, five main issues have been identified and are discussed under the headings; the context in which the research was set, the research design, the preliminary study, the design of the main study, and the data analysis plan.

### Context

The research was to be carried out within an ongoing Mathematics Education course which was part of the Bachelor of Teaching (primary) degree at the University of New England. In order to describe the context for the research strategy, a description is given of the university and of the course.

The University of New England is situated in a rural area of northern New South Wales (NSW). Consequently, many of the students who attend the University come from predominantly rural backgrounds. In addition, a significant proportion of the students have undertaken their secondary education in Queensland as well as New South Wales. The University is unique in that there is provision for most students to reside on campus. The University has four faculties, namely, Arts; Economics, Business and Law; Education, Health and Professional Studies; and The Sciences.

The amalgamation of the Armidale College of Education with the University of New England in 1989 lead to the conversion of the three-year Diploma in Primary Education to the Bachelor of Teaching degree. This degree aims to cater for the professional and personal development of students who are about to enter the teaching profession. The course is set within the Australian context, and is based on the fundamental premise that anyone wishing to become a teacher must acquire a thorough understanding of how children learn. The course contains three strands, education studies, general studies and curriculum studies. The

two Mathematics Education units are part of the curriculum studies strand which relates to the content and teaching of the primary school curriculum

The course titled Mathematics Education I is the first of two compulsory Mathematics Education units. It is undertaken by all primary-teacher education students during the second semester of their first year of candidature as part of the Bachelor of Teaching degree. A description of the course, given below, was detailed in the University of New England *Handbook* (1992, p.503).

> The unit aims to consolidate the mathematical skills of the student, as well as to develop an understanding of teaching mathematics K-6. The unit covers the three content areas of the NSW Mathematics Syllabus (1989), i.e., number, space and measurement, together with teaching inquiry skills of particular importance to mathematics. The unit also includes a component designed to determine students' basic mathematical competencies with appropriate remediation if necessary.

The course structure allows for five hours of tuition each week for one semester. This consists of three one-hour sessions, a lecture, a student-prepared seminar, and a basic mathematical competency tutorial, and a two-hour workshop during which the students explore hands-on teaching materials and develop teaching enquiry skills. The unit content includes a section on the understanding and teaching of 2-D and 3-D space, as well as a component designed to determine the students' basic geometric competencies. These aspects are addressed in the latter part of the course, i.e., during the last two weeks.

**Limitations and constraints**

In planning the research design, account had to be taken of limitations and constraints imposed by the course structure. The willingness of students to participate in a geometry test and the time they were prepared to allocate to the research needed to be considered. To guarantee an adequately large number of students who would be committed to give their best, the decision was taken to include the test in the course structure as part of the basic mathematics competency segment, the understanding of geometry.

This decision meant that only two hours of class time were available, one hour to do the paper and one hour for feedback and discussion. The disadvantage of this decision was that it appeared that it might not be possible to test all students in all concepts in a one-hour time interval. The advantages were that students saw the test as part of the program, hence there was incentive to be present and to participate fully.

By limiting the test to one hour, the risk of exceeding the students' concentration span would be minimised. Additionally, with the possibility of interviews following, students may consider a two-hour test too demanding of their time. Further, the purpose of the investigation was not in the collection of large amounts of data. It was to get reliable data across a number of topics and questions, i.e., to have students give their best. The focus of the thesis was an intensive consideration of the meanings or understandings students apply to the test items.

## Research Design

In designing the research, several issues needed to be considered. These included the structure of the preliminary study, the van Hiele levels and geometric concepts which were to be included, time limitations, and changes necessary to the interview format used by Mayberry.

### Design of the preliminary study

The purpose of the preliminary study was to develop an appropriate written test. Mayberry (1981, p.101) recognised the usefulness of a paper-and-pencil test, querying whether such a test "could be developed to identify a student's level." She focused solely on the multiple-choice test, commenting that it "would be very difficult to develop and to analyze." The decision to carry out the investigation of the Mayberry items in an Australian context as a written test rather than in an interview situation resulted from the following considerations. Clinical interviews, while giving an in-depth knowledge about a subject are very time consuming, hence they are not particularly suitable for assessing large numbers of students. The alternative, testing a small number of students can limit the generalisability of the findings of a research study. For example, since each interview in Mayberry's study had a duration of approximately two hours, Mayberry was able to test only nineteen students.

A second consideration concerns the prospect of using the test to assess the van Hiele level of school students. The time available in a classroom for assessment is limited. Should the test be adaptable, in part or whole, for determining the level of understanding of schools students, a written test could be more appropriate than an interview schedule.

In translating the interview items designed by Mayberry into a written test format suitable for use in Australia, four main questions were considered:

1. Should all five van Hiele levels be included in the test?
2. Are the seven concepts tested by Mayberry appropriate in an Australian context?
3. Can all the concepts be tested in the one-hour time limit?
4. Does the phrasing in the interview questions need adjusting to ensure
    (a) clarity in the written format, and
    (b) use of expressions recognisable in an Australian context?

*Should all five van Hiele levels be included?*

The main decisions to be considered concern Levels 1 and 5. The five van Hiele levels examine students' understanding of a range of skills, from their visual recognition of a geometric concept to their abstract appreciation of rigor. The students to be tested ranged in geometric background from those who had studied little or no geometry at secondary level to those who had completed a formal geometry course in senior secondary school. Mayberry's subjects were drawn from a similar cross-section of students. The results of her study showed that some of the students failed to demonstrate recognition (Level 1) of one or more of the concepts. Because of this, and because of the comparability of the backgrounds of the two groups of subjects, it was considered necessary to determine whether all students were able to achieve visual recognition of the concepts. Hence, questions testing for understanding at the first level were included.

When deciding whether to include test items for Level 5, four questions were considered. First, does Level 5 exist and if so, is it testable? Second, what were Mayberry's views on the testing of Level 5? Third, did any of Mayberry's subjects demonstrate Level 5 thinking, and finally, was the sample likely to contain students working at Level 5?

In his study *Van Hiele Levels and Achievement in Secondary School Geometry*, Usiskin (1982, pp.77-79) found that although the level theory is an elegant theory regarding the acquisition of an understanding of geometry as a mathematical system, there was confusion in the Chicago project in interpreting the van Hiele statements concerning Level 5. This led to difficulty formulating suitable Level 5 test questions. The confusion, Usiskin considered, was supported by Pierre van Hiele's disavowal of Level 5 in his more recent writings. Usiskin (p.79) concluded: "In the form given by the van Hieles, Level 5 either does not exist or is not testable. All other levels are testable."

Mayberry (1981, p.51), while including Level 5 in her test, considered questions at this level required different treatment to questions testing for Levels 1 to 4. She noted "these questions

are referent free and, therefore cannot be put into concept categories." None of Mayberry's subjects were successful with the Level 5 items.

Finally, the geometric background of the students in this study needed to be considered. Although several of the students had studied formal geometry in senior secondary school in both NSW and in Queensland, the focus of their work was only early Level 4. As a result, it was considered unlikely any of the subjects would have been in an environment to consider Level 5 issues, or attained Level 5 reasoning. This view of secondary students' thinking is supported by van Hiele (1986, p.47), "In school we have to deal with Levels 2, 3, and 4. If our pupils do not understand us, it is at these levels, not when we are speaking on the fifth or perhaps still higher levels."

For these reasons it was decided to design the test to include questions for Levels 1 to 4, but to exclude Level 5. This meant that the test would include the first fifty seven of Mayberry's items but omit the Level 5 items, i.e., items 58 to 62.

*Are the seven concepts tested by Mayberry appropriate in an Australian context?*
Mayberry, in her study, selected seven topics, namely, square, right triangle, isosceles triangle, circle, parallel lines, congruency and similarity. The topics were chosen because they were in the elementary curriculum and were included in elementary textbooks. Mayberry (1981, pp.49-51) assumed that her subjects, the college students, had been exposed to textbooks containing the chosen concepts. The inclusion of the topics in the curriculum inferred that they "therefore seemed essential to the geometric understanding of the potential teacher" (p.49). All these concepts are relevant to both the New South Wales and the Queensland secondary syllabuses. Hence, the decision was taken to include all seven concepts in the test, despite time limitations.

*Can all the concepts be tested in a one-hour time limit?*
One of the constraints affecting the Mayberry research was the small sample size necessitated by the interview technique. The resultant limiting of the generalisability of her results led to the formulation of one of the objectives for this research, to test a significantly large sample of students.

As detailed above in the limitations and constraints, the test itself was allocated a one-hour period in the Mathematics Education I course In her study, Mayberry's interviews lasted approximately two hours. It was estimated that the time required to complete the fifty-seven items for Levels 1 to 4 in a written format would approach two hours. This led to the

decision to design two test papers of similar format which, between them, covered the seven concepts. The papers were each to include questions on four concepts. The concept square was considered to be the most familiar of the seven concepts and, therefore, the items relating to this were included in both papers. One of each of the two triangle types selected for inclusion by Mayberry, right triangle and isosceles triangle, was allocated to each paper. The remaining four concepts were divided evenly between the two papers. This resulted in Paper I containing the items testing the concepts square, right triangle, circle and congruency, and Paper II containing the items testing the concepts square, isosceles triangle, parallel lines and similarity. Mayberry did not design an equal number of items per concept. Hence, the resultant Paper I for the preliminary study contained thirty-four items and Paper II, thirty-seven items. The items were not renumbered in the written papers, but left with their original Mayberry numbering to facilitate comparison in the assessment of the responses and the results.

*Does the phrasing in the interview questions need adjusting?*

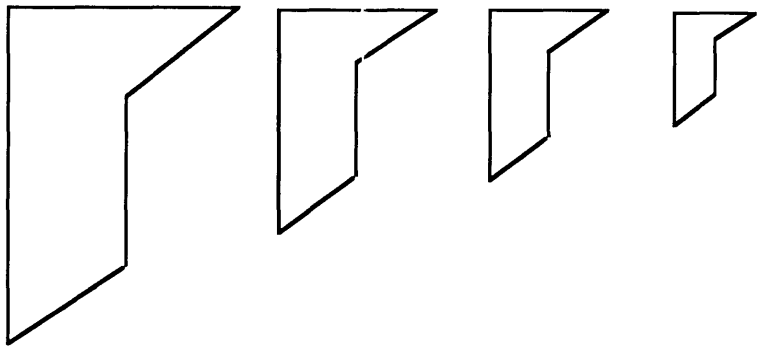
Two factors needed to be considered in ensuring clarity of the written questions. An interview situation allows for the rephrasing of a question and for the inclusion of additional instruction should there be any confusion on the part of the subject. Therefore, some of the Mayberry interview items needed modifying to facilitate clarity in the written paper. Second, any specifically American symbols and expressions needed replacing with a format recognisable by Australian students. This led to the following adjustments to format and phraseology in the Mayberry interview items:

1. When a question required a simple yes/no response, YES—NO 'boxes' were inserted.

2. Space was provided for written responses. When an item required more than one answer, the individual questions were separated and space provided for each answer.

These two adjustments are illustrated in alterations to the wording of Item 8:

Item 8



Question (Mayberry)
Are these figures alike in any way? What word describes this?

Question (Australian)
Are these figures alike in any way? YES NO
What word describes this? _____ _____

3. Symbols and expressions were changed to those which are more commonly recognised in Australia. For example, 'interval OB' became simply 'OB', and the American congruency sign '≅', depending on the context, became either the equal sign '=' or the congruency sign '≡' which is most commonly used in Australia.

4. Wording was changed to facilitate clarity, e.g., in Item 31, 'Which of the following is/are true?' became 'Are the following true or false?'

5. Oral prompts, which were suggested for the Mayberry interviews, were incorporated in the written questions. For example, Item 17 asks "Does a right triangle always have a longest side?" The Mayberry interviews included the following prompt 'If student says yes, ask which one?' In the Australian paper the item became:

Item 17 (Australian)

Does a right triangle always have a longest side? _____

If so, which one? _____

6. Items 24, 26, 28, and 39 gave a list of necessary and sufficient conditions for the concepts square, right triangle, isosceles triangle and parallel lines, respectively. Each Mayberry question asked "Which combination of the following guarantees a figure to be a (square)?" Should the response contain too many conditions, the interviewer was

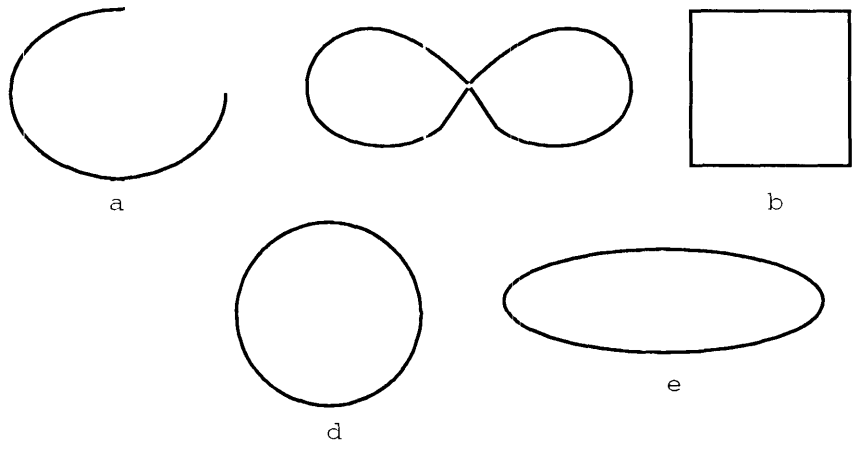
instructed to probe twice 'Can you use less?' The instructions for each item in the written papers were amended to the format given for Item 24.

Item 24 (Australian)

| |
|---|
| Circle the smallest combination of the following which guarantees a figure to be a square. |

7. The Mayberry test requires that Item 33 be assessed at two levels, properties of a circle (Level 2) and definition of a circle (Level 3). Should a response fail to indicate Level 3 reasoning, an interviewer is able to probe more deeply. This is not possible in a written paper. The need for a Level 3 response was indicated to the respondents in the written paper by the inclusion of an additional question.

Item 33




Question (Mayberry)
Tell why each of the figures is or is not a circle.

Questions (Australian)
Tell why each of the figures is or is not a circle. (Spaces (a) to (e) were provided).

Can you give a general rule to fit all the above answers? _____

8. During the Mayberry interviews, particularly following the completion of Item 44, the terms congruent and similar were explained to the Mayberry subjects. This was because many students were found to be using the word equal incorrectly. It was decided that since congruency and similarity are part of the junior secondary curriculum in Australian schools, this information should be unnecessary for the written test, and, hence, was omitted. Additionally, it was considered of value to ascertain whether or not these terms were understood.

Details of the preliminary test papers can be found in Appendix A.

## The Preliminary Study

The following description of the preliminary study includes details of the administration of the test and of the subjects, discussion on the evaluation of the responses, scoring of the responses, recording of the results, the interviews, and details of further changes for the main study.

### Administration of the preliminary test

Thirteen university students volunteered to participate in the preliminary study. The students were all residents of one of the campus colleges, and included both undergraduates and post-graduates, students recently from secondary school, and mature-age students. Their mathematical backgrounds ranged from students with no senior secondary mathematics to students with university majors in mathematics. The papers were allocated randomly to the students. Six students completed Paper I and seven students completed Paper II. No time limit was set for the tests. Students took between forty-five and sixty-five minutes to complete the papers. The least time was taken by the more competent students and by those with very little mathematical background. The average completion time for Paper II was slightly longer than for Paper I, possibly resulting from the greater number of items in Paper II.

### Evaluation of the responses

If the results of the study were to be comparable with the Mayberry results, the evaluation of the responses needed to replicate Mayberry's evaluation standards as closely as possible. A detailed examination of the Mayberry thesis led to the identification of two factors, namely, that the response to each question is correct only if it reflects the level of thinking for which the question has been designed, and that the response should correspond with Mayberry's behavioural descriptions of the van Hiele levels and her expectations of suitable responses.

First, Mayberry designed each question to test a specific van Hiele level. A response was credited as correct only if it reflected the designated level. The researcher, therefore, needed to be aware of the characteristics Mayberry (1981, p.10) considered a student would display if geometric thought conformed to the van Hiele theory were noted. If a question could be answered at more than one level, the student would respond at her/his present level of reasoning. A student working at a level below the highest at which a question could be answered correctly would fail to be fully aware of the intent of the question. Should the question be able to be answered at only a single level, a student, unable to perform at that

level, would fail to be aware of the intent of the question, hence would fail to answer the question correctly or fail to answer the question at all. For example, should a student's working level be Level 2, identification of properties, and the question be testing for knowledge of the relationships between properties (Level 3), the student's answer would most likely be a simple listing of known properties. The student would neither be likely to comprehend that there is a relationship between the properties, nor be aware that this is what the question wanted as an answer. Lack of comprehension of a Level 3 /4 question may be indicated by comments, such as "You have omitted the diagram" or "I don't understand what you are asking." Responses, such as these, would not be credited as correct for the Level 3 question.

Second, if there was to be a comparative evaluation, the researcher needed to be aware of the nature of the responses which Mayberry considered acceptable for each question. Not only were Mayberry's behavioural descriptions of the van Hiele levels (pp.47-49) analysed, but also the comments made by the expert judges on the initial set of questions (pp.53-57). Mayberry's analysis of the responses of her subjects to particular questions and clusters of questions (pp.78-85), together with her comments on these responses were also taken into consideration. From this detailed analysis an evaluation procedure was developed.

**Scoring of the preliminary test**

In her study, Mayberry developed a formula for assessing the responses to her items. First, Mayberry designed each item to be consistent with the behavioural terms of a specific van Hiele level, i.e., each item measures a student's ability to reason at that level. Second, some items contained more than one question part, hence, Mayberry weighted each question part. If the responses to the parts met the weighting for that question, one mark was scored. Correspondingly, if the weighting was not reached, the response(s) scored a zero. Third, Mayberry set a critical score for each level within each concept. The total score for the responses for the level was then compared to Mayberry's critical score. A student achieving the criterion or better for a level was deemed to be reasoning at that van Hiele level. For example, Item 9(b) was designed to test the Level 3 skill, class inclusion, for the concept square. The item contains four question parts. Three correct responses (out of four) earns one mark towards the Level 3 score for the concept square. Of a possible maximum score of nine for this level, the criterion was set at six, i.e., any student scoring 6, 7, 8 or 9 would be assessed as successful for van Hiele Level 3 reasoning, for the concept square. A student scoring five or less was deemed not to be reasoning at that level. A student's response to a question was not assessed as an indicator for any other level other than the level for which the question was designed.

The results for each concept were then recorded as an ordered quadruple. A "1" was given if a subject attained the critical score for a level, and a "0" if the criterion was not met. For example, a quadruple (1,1,1,0) indicates that a subject has met performance criteria for van Hiele Levels 1, 2 and 3, but not for Level 4. Thus for each concept, with two possibilities for each level component, there are $2^4$ or sixteen possible response patterns for each quadruple. Since each paper contained questions for four concepts, each subject's results would consist of a set of four quadruples, one for each concept. For both the preliminary and main studies in this research, assessment formats based on Mayberry's tables of scoring criteria were drawn up (see Appendix B). If the van Hiele levels are hierarchical, as shown by Mayberry (1981) and Denis (1987), then a student demonstrating reasoning on Level $n$ should also have performed on Level $n$-1. Hence, of the sixteen possible response pattern errors, only five ordered quadruples should appear, (0,0,0,0), (1,0,0,0), (1,1,0,0,) (1,1,1,0,) and (1,1,1,1).

Non-scale patterns, i.e., those which are not one of the five patterns showing hierarchy are deemed to contain response pattern errors. Such an error is not a wrong answer, but a response which could have led to an incorrect prediction of a student's van Hiele level of reasoning if based on the subject's highest level. This occurs when a student has failed to reach criterion at one level, yet has achieved success with a higher level. For example, a student may have been successful with Levels 1, 2 and 4, but failed to reach criterion for Level 3. This would be recorded in an ordered quadruple as (1,1,0,1).

**Results for the preliminary study**

The responses of each student were assessed according to Mayberry's scaling. Full results together with the quadruples are tabled in Appendix C. Overall results for the preliminary study showed that the written tests were scoreable. Tables 3-1 (Paper I) and 3-2 (Paper II) summarise the results of the thirteen students (coded as P01 to P13), showing the highest level achieved by each student for each concept. In assessing the results, it was found that due to the structure of the Mayberry test questions and assessment method, it was not always possible to allocate a value for Level 4 in the concept similarity. This difficulty and the proposed solution is detailed in the changes made to the written test for the main study.

Table 3-1

Highest Level Achieved by Students in Preliminary Study (Paper I)

| Student | Square | Right Triangle | Circle | Congruency |
|---|---|---|---|---|
| P01 | 4 | 4 | 4 | 4 |
| P06 | 4 | 3 | 4 | 4 |
| P08 | 4 | 3 | 4 | 4 |
| P09 | 2 | 2 | 3 | 2 |
| P11 | 2 | 2 | 3 | 3* |
| P12 | 3 | 1 | 3 | 4* |

* response pattern error

Table 3-2

Highest Level Achieved by Students in Preliminary Study (Paper II)

| Student | Square | Isosceles Triangle | Parallel Lines | Similarity |
|---|---|---|---|---|
| P02 | 2 | 4* | 3 | 3# |
| P03 | 2 | 3* | 3 | 2# |
| P04 | 2 | 2 | 2 | 2# |
| P05 | 3 | 3 | 2 | 2# |
| P07 | 4 | 4 | 4 | 3# |
| P10 | 3 | 3 | 2 | 3# |
| P13 | 3 | 3 | 3 | 3 |

* response pattern error # Level 4 was not able to be assessed

On four occasions students failed to show mastery at Level $n$-1 when successful with Level $n$ items. These occasions (referred to as response pattern errors) are indicated by an asterisk in Tables 3-1 and 3-2. The errors occurred in two concepts, isosceles triangle and congruency, with two errors in each concept. Three of the errors occurred between Levels 2 and 3, while the remaining error occurred between Levels 3 and 4. The two students showing response pattern errors in isosceles triangles, P02 and P03 were both interviewed, and details of the errors are given in the next section. P11 and P12 were unavailable for interview. The details resulting in their incurring an error in their response patterns are given below.

P11 reached criteria for Level 1 (2 out of 2) and Level 3 (7 out of 10, criterion of 6) in the concept congruency, but only gave two correct answers to Item 22's four question parts (criterion 3 out of 4) for Level 2. She identified correctly the corresponding sides and angles, yet appeared to misinterpret the descriptive questions. Her responses to Q22 (b and d) were:

Q22(b) What is true about their sides?
P11 *There are four of varying length.*

Q22(d) What is true about their angles?
P11 *Add up to 360°*

In her responses to the Level 3 questions for congruency, P11 demonstrated partial knowledge of relations between properties and their implications, but no awareness of class inclusion. Overall, P11 showed comprehensive knowledge of properties of the concepts, but only partial awareness of their relationships, only reaching Level 3 criteria for the concepts circle and congruency.

A response pattern error was registered by P12 between Levels 3 and 4, in the concept congruency. In her responses, P12 demonstrated confidence in expressing her knowledge of properties of concepts, and the relationships between the properties for all four concepts in Paper II. However, she did not demonstrate the same degree of confidence when challenged with questions concerning implications of the properties, and class inclusion. This is shown in her responses to the following question parts for Item 43, in which she considers the possibility of equality for pairs of corresponding sides (part a) and non-corresponding sides (part d) in congruent triangles. Her responses show an ability to identify corresponding sides, and awareness that corresponding sides are always equal, but appears to imply from this, that non-corresponding sides cannot be equal. She seems unable to consider the possibility that the triangles may be isosceles, and, hence, that this condition would allow equality of non-corresponding sides in a pair of congruent triangles.

Item 43

| |
|---|
| Δ ABC is congruent to Δ DEF (in that order).<br>Are the following a) certain b) possible, or c) impossible?<br>Give reasons for your answers. |

a) AB = DE

P12 *Certain, corresponding sides in congruent triangles are the same.*

d) AB = EF

P12 *Impossible, sides don't correspond.*

This limitation in Level 3 understanding resulted in her responding correctly to some Level 3 questions in all concepts, but meeting criteria only for the concepts square and circle. For the concept congruency, P12 scored 5 out of 10, one short of the Level 3 criterion, yet she demonstrated skill in formally proving congruency of triangles for the Level 4 question, Item 56, thus meeting the criterion of 1 correct response out of 2. This led to her registering the response pattern error.

Analysis of the responses for P11 and P12 indicated the response pattern errors appeared to be due to mis-interpreting the intention of a Level 2 question (P11), and to having incomplete mastery of Level 3 while able to respond correctly to an easier Level 4 question (P12). While these errors would be considered when evaluating the responses in the main study, they were not taken as an indication of need to modify the test.

**Interviews of preliminary study subjects**

Following the scoring of the written tests, eight of the subjects from the preliminary study were interviewed. The remaining five students were unavailable due either to illness or to other personal commitments. The focus of the interviews was to ascertain:

a) whether the test questions were understood clearly, and whether any expressions cause problems; and

b) whether the levels displayed by the students in the written test were stable.

The interview format was unstructured, allowing the interviewer freedom to prompt and probe students to ascertain (a) whether further clarification of the questions was needed, and (b) whether the students were able to give responses at a higher level than that those in the written paper. The interview questions, therefore, were designed individually for each student to determine the reason for their various errors and omissions. This meant that thirty-two of the fifty-two (four concepts for each of the thirteen students) assessments of levels in the written paper were able to be checked for validity.

The interviews confirmed twenty-nine of the level assessments. Of the three levels not confirmed, two related to the Level 2 response pattern errors for the concept isosceles

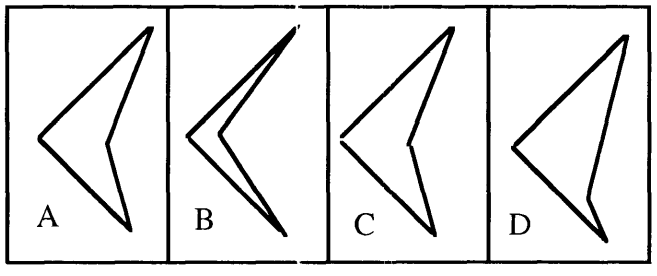
triangle. In each case, the error appeared to have been due to a misreading of the question. With probing, both students indicated they had read the easier questions too quickly, and, hence, not perceived the intention of the question. There was no indication in either case that the misinterpretation resulted from a lack of clarity in the question. Neither student hesitated in giving a correct response at the interview. The third occasion of a level not being confirmed occurred with a subject who was a mature-age language student who had not encountered geometry since completing secondary education many years earlier. This student, P03, found that the probing in the interview caused recollection of several facts, particularly in the concept square in which she then demonstrated a higher degree (namely Level 3) of understanding.

This confirmation by interview of twenty-nine out of thirty-two (91%) level assessments in the preliminary tests was considered to be an indicator that the written format of the Mayberry items was acceptable. As anticipated, all students interviewed, demonstrated an appropriate awareness of the terms congruency and similarity. This enabled them to proceed adequately on the relevant questions. However, the interviews did reveal two weaknesses. First, the phrasing of Item 15 was not clear, and, second, questions did not test consistently the concept similarity for Level 4 understanding. Details of the alterations made to these questions are described below.

**Changes made in the preliminary paper for the main study**

Analysis of the test results and the ensuing interviews indicated that two alterations were necessary. First, the interviews revealed that there was a problem in the clarity of Item 15 which had led to some confusion.

Item 15



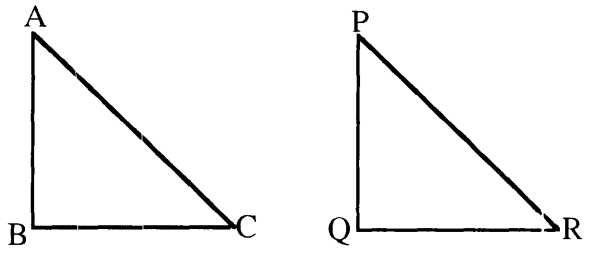

Question (preliminary study)
Which figure **is** congruent to A? _____

Question (main study)
Which figure **appears to be** congruent to A? _____

There was doubt that the diagram showed two identical figures clearly as inferred in the question. This problem was remedied by changing the wording to that used by Mayberry in similar items, namely, Items 10, 11, 13 and 14. The alteration to the wording is shown above in bold type.

The second change, indicated by the interviews, resulted from an omission which may not have arisen in an interview situation. Item 48 was designed by Mayberry to test for Level 4 understanding in three concepts, right triangles, isosceles triangles and similarity. Significantly, it was the only question designed to test Level 4 for the concept similarity. Because of the distribution of the three concepts between the two written papers, the item occurs in both Papers I and II, hence is attempted by all students.

Item 48




Figures ABC and PQR are right isosceles triangles with angles B and Q being right angles. Prove that $\angle A = \angle P$ and $\angle C = \angle R$.

Two commonly presented and correct solutions to this question involve the use of either the properties of similarity, or the angle properties of isosceles and right triangles. In the preliminary study, only two students (one for each test paper) used similarity in their proof, the other eleven preferring to choose angle properties. This resulted in only one of the seven students, attempting Paper II, showing whether he/she could demonstrate Level 4 reasoning for the concept similarity. There was no other item which could indicate Level 4 reasoning in similarity in the paper for the remaining six students. To overcome this difficulty of not necessarily being able to assess whether a student has Level 4 understanding of similarity, a new question, Item 58 was created. The item incorporates the Level 4 concept of understanding the meaning of necessary and sufficient conditions.

Item 58

Figures MNO and PQR are right triangles with $\angle N$ and $\angle Q$ being right angles. MO and PR are in the ratio a:b. What is the least additional information needed to ensure that triangles MNO and PQR are similar?

The concept similarity is tested in the second written paper. The addition of Item 58 resulted in the composition of the papers for the main study being thirty-four items in Paper I and thirty-eight items in Paper II. The intention of replicating the Mayberry test as closely as possible meant that the imbalance in length between the papers could not be avoided. The final format of the written questions for the main study is shown as changes to the preliminary test questions in Appendix A.

**Design of the Main Study**

The purpose of the main study was to investigate four research themes:

1. How does a sample of Australian students training to be primary teachers perform on a written test version of the Mayberry questions?

2. Can a quantitative analysis of the results, using a partial credit model, offer further insight into the nature of the Mayberry test?

3. Do the Mayberry items measure the van Hiele levels for which they have been designed? Are the success criteria, established by Mayberry, valid?

4. What implications can be drawn when an alternative assessment system, such as the method used by Gutiérrez, Jaime and Fortuny, which does not have the discreteness of levels as a significant feature to influence the findings, is applied?

The preliminary study had resulted in the development of a set of good written test items divided between two papers. For the main study, students enrolled in the Mathematics Education I course (1992) were allocated randomly to one of the papers. This meant that approximately half of the students would answer questions on the concepts, square, right triangle, circle and congruency, and the remainder would answer questions on the concepts, square, isosceles triangle, parallel lines and similarity.

The tests were administered to the students during a one-hour period in the third week of the semester, and this was followed by a one-hour feedback and discussion time. Before they commenced the written test, the students were given the following instruction both in writing and orally:

> This test is much more than seeing whether you are right or wrong. In fact it is more important to us to know how you think during a question, and the processes you are using to find an actual answer (although we would like you to get as many right as possible). Therefore could you take care in presenting as much detail as possible when the question asks for you to give REASONS.

Prior to the test, the students were not informed of the topic. On finding that the subject of the test was geometry, there was perceptible dismay among the students indicating a degree of anxiety concerning the subject. Following assessment of the written responses, ten percent of the students were interviewed to provide additional information of the levels displayed in the written test. The testing and interviewing were completed before either of the geometry sections in the Mathematics Education 1 course were commenced.

In designing the main study, the following points were considered:

1. the composition of the sample;
2. selection of the interviewees;
3. structure of the interviews; and
4. the data analysis plan.

**The subjects**

Sixty-one (coded S01 to S61) of the sixty-eight students enrolled in the Mathematics Education I course (1992) completed the test papers, thirty one completing Paper 1 and thirty completing Paper 2. Fifty three of the students were female, and eight were male. All students were asked to complete an information sheet detailing their mathematical background. The completion of this information sheet was optional in compliance with the privacy clause. Fifty-five students volunteered information about their secondary mathematical background. Of the fifty-five students, twenty nine had completed their secondary schooling in New South Wales, twenty four in Queensland, one in Victoria and one in the United States of America. Their mathematical backgrounds were summarised into three groups based on the geometric content of their secondary mathematics courses. The first group consisted of those students who had completed a senior secondary mathematics course which included a formal or recognised geometry segment (minimum level — NSW, 2-unit Mathematics; Queensland, Mathematics-I). The second group consisted of students who had completed a senior secondary mathematics course which did not contain a formal geometry segment (NSW, Mathematics-in-Society; Queensland, General Mathematics). The third group consisted of students who had not completed any senior secondary mathematics. Thirty five of the students were in group 1, thirteen in group 2, and seven in group 3. Table

3-3 summarises the students' mathematical background together with the location of their senior secondary schooling.

Table 3-3
Mathematical Background and School Location for Subjects (n=55)

| Geometric background | Total | NSW | Queensland | Other |
|---|---|---|---|---|
| Senior geometry | 35 | 23 | 12 | 0 |
| Senior maths but no geometry | 13 | 4 | 7 | 2 |
| No senior maths | 7 | 2 | 5 | 0 |

**The interviews**

Six students were interviewed. How the interviewees were selected, and the format of the interviews is now considered.

*Selection of interviewees*

The students were graded into three strata according to their most common level of working, as indicated by the results. The strata included as broad a representation of student thinking as possible. The results showed that a majority of the students were giving Level 2 responses. This led to the three strata being composed of students whose responses generally displayed the thinking of the following levels;

a) early to mid-Level 2 (to be referred to as Level $2^{o)}$

b) confident Level 2 together with some early Level 3 (Level $2^{+)}$

c) confident Level 3 together with some early Level 4 (Level $3^{+)}$.

The six students represented equally the three strata, the two genders and both written papers. When considering to which stratum each student belonged, the thrust of their written responses was taken into account as well as their results as determined by the Mayberry criteria. It was considered that this would allow for a more accurate allocation of each student to a stratum since the Mayberry allocation of levels to items and the setting of criteria were yet to be evaluated. Table 3-4 shows the students selected for interview, together with a quadruple showing the highest van Hiele level reached by each student for each concept tested in the written paper. It should be noted (see asterisk in Table 3-4) that S08's responses indicated his level of thinking was mid-Level 2. His responses did not give any indication of ability to relate properties; he was considered to be the student who best fitted the interview category. His Mayberry Level 3 rating for the concept circle was found to be an incorrect

assignment of level, resulting from a feature of the Mayberry testing which is discussed in Chapter 6.

Table 3-4
Students Selected for Interview in the Main Study

| Stratum | Male students | Female students |
|---|---|---|
| Level $2^{o}$ | S08<br>Paper I<br>(2,2,3*,2) | S59<br>Paper II<br>(2,0,2,1) |
| Level $2^{+}$ | S31<br>Paper I<br>(2,2,2,2) | S33<br>Paper II<br>(2,3,2,1) |
| Level $3^{+}$ | S41<br>Paper II<br>(4,4,2,3) | S14<br>Paper I<br>(4,3,4,4) |

* See comments above concerning this result.

*The interview format*

Eight Mayberry items were selected for the interviews, two for the concept square and one for each of the other six concepts. This meant that five items occurred in each of the original written test papers. The items selected were those among the Mayberry items, considered most suitable for probing for different van Hiele levels of response. Table 3-5 shows the items which were selected, together with the concept and levels of each.

Table 3-5
Characteristics of Items Selected for the Interviews

| Item | Concept | Mayberry Level(s) | Probing Levels |
|---|---|---|---|
| 23 | square | 2 and 3 | 2 and 3 |
| 27 | right triangle | 3 | 2 and 3 |
| 35 | circle | 3 | 2 and 3 |
| 39 | parallel lines | 3 | 2, 3 and 4 |
| 41 | similarity | 3 | 2 and 3 |
| 45 | square | 4 | 2, 3 and 4 |
| 50 | isosceles triangle | 4 | 2, 3 and 4 |
| 55 | congruency | 4 | 2, 3 and 4 |

A response to a question may or may not show a student's correct level of thinking. This problem can occur for several reasons. The question may have been misread, or part or all of it may have been misinterpreted. The format of a written paper may have been restrictive, not allowing students to express themselves fully. There may be other possibilities. A well-designed interview schedule should permit clarification of the overall picture of how the student is thinking. If an error has occurred in the evaluation of a student's thinking, the schedule should allow identification of the cause of the error through clarifying the components of the problem. The interview schedule for this study was based on the error-analysis approach detailed by Newman (1977), and designed to clarify three points:

a) Were the items understood?
b) Did the level of the interview response correspond with the level given in the written answer?
c) Did the students change the level of their response when challenged?

Each subject was interviewed individually by the experimenter on the set of interview questions (see Appendix D). Each interview lasted approximately thirty minutes, and each subject was asked to respond to all eight items having attempted previously five of the items, but not seen the other three. It was considered the responses to the unseen items should reinforce the overall working level of the subjects as indicated by their total responses. The subjects were given a copy of each item in the same format as in the original written test papers. They were not given the prompting and probing questions. Each item had space for the subjects to write down their response which complemented or replace the verbal response. All interviews were audio-taped. In addition, the interviewer had recorded on paper any impressions not otherwise noted. On completion of all six interviews the recordings were transcribed.

**Data Analysis Plan**

The evaluation of the study is considered in terms of the four research themes, i.e., the study needs to provide answers to the following questions:

1. a) At what van Hiele levels did the students' respond overall, and for each concept?
   b) How did the levels demonstrated by the students compare with their senior secondary mathematical background?
   c) How did the students' responses compare with Mayberry's observations?

2. What insight into the Mayberry test is offered by a quantitative analysis of the results, using a partial credit model?

3. a) Do the items measure the level for which they were designed?
   b) Are Mayberry's success criteria valid?

4. Can the items be evaluated using the Gutiérrez, Jaime and Fortuny assessment scheme, and how do the results compare with Mayberry's marking scheme?

To provide the required information, the responses for the main study were analysed in the following manner. First, the responses were evaluated using the Mayberry standards developed for the preliminary study. The total of correct responses was then compared to Mayberry's criterion for each of the four concepts and four levels tested in each paper. These results were recorded as four sets of ordered quadruples. Second, the scalability of the results was verified by applying the Guttman scalogram analysis. Third, the overall results were compared with the students' senior secondary mathematical backgrounds, and fourth, the results of the Australian students were compared to the Mayberry results. Fifth, the results would be analysed using the QUEST program (Adams & Khoo 1993) which involves the utilisation of the Rasch measurement theory to cases where multiple response categories are provided. Finally, the responses would be reassessed according to the Gutiérrez *et al* method, leading to a comparison between the two evaluation methods.

**Methods of analysis**

Three techniques were used in the analysis of the results, the Guttman coefficient of reproducibility, the Rasch measurement theory, and the assessment method developed by Gutiérrez, Jaime and Fortuny (1991). A description of each follows.

The Guttman coefficient of reproducibility compares the total number of errors occurring as shown in non-scale patterns of results to the total number of responses. As detailed in the section describing the scoring method for the preliminary study, each student's results would be expressed as a set of four ordered quadruples. Of the sixteen possible response patterns for each quadruple, all except the five hierarchical patterns contain one or more response pattern errors. The number of errors in an ordered pattern is defined as the smallest number of changes necessary to make the pattern conform with one of the acceptable patterns. For example, the quadruple (1,1,0,1) contains one error while (1,0,0,1) contains two errors.

The coefficient of reproducibility (Rep), i.e., the measure of error for the entire scale is given by the formula:

Rep = 1 - total number of errors/ total number of responses

In this study, the total number of responses was 4 (levels) * 4 (concepts) * 61 students. Hence

Rep = 1 - total number of errors / 4*4*61

Torgenson (1967, p.323) stated:

> Rep is the primary criterion of scalability. A Rep of 0.90 or better has been taken as the standard. A value of Rep equal to 0.90 or more means that, of all of the responses of all of the subjects to all of the items, no more than ten percent correspond to errors of reproducibility.

Mayberry used the Guttman scalogram analysis (the calculation of the coefficient of reproducibility) to confirm that the van Hiele levels, as tested by her items, were hierarchical in nature. If the Mayberry test items are translated successfully from an interview situation to a written paper, i.e., the translation does not affect the essence of the Mayberry test, then the value of the coefficient of reproducibility for the Australian results should be similar to the value calculated for the Mayberry results. In Mayberry's study, the ordered patterns for the nineteen subjects contained twenty-four errors which, over the first four levels, gave a coefficient of reproducibility of 0.95 or 95%. The Guttman scalogram analysis for this research is presented in Chapter 4.

Combined data analysis, the results of which are presented in Chapter 5, was used to substantiate many of the general observations made in Chapter 4. The QUEST program (Adams & Khoo 1993) drew a comparison between the relative degrees of difficulty of the Mayberry question parts, demonstrated patterns which support the level structure, and produced an estimate of geometric understanding for each student. It incorporated the information from the responses to the questions covering all the concepts for Papers I and II. The QUEST program is an implementation of the Rasch latent trait scaling model. This was first introduced for the analysis of dichotomously-scored responses in 1960. When data are fit to the model, person parameters can be freed from the item difficulties and item parameters can be estimated independently of the calibrating sample. Since being first introduced, the Rasch model has been extended to analyse polychotomous data such as counts, repeat trials and rating scales.

The partial credit model is relevant to items where responses can be coded according to an increasing or decreasing degree of 'correctness'. Rather than apply the Rasch dichotomous model to each step in performance level, which assumes hierarchical independence and increasing difficulty with each step, Masters (1982, pp.157-158) suggested considering the individual difficulty of each successive step in the item. The probability of student n responding in category x of item i is given by

$$p_{nix} = \frac{\exp \sum_{j=1}^{x} (\beta_n - \delta_i - \tau_j)}{1 + \sum_{k=1}^{m} \exp \sum_{j=1}^{k} (\beta_n - \delta_i - \tau_j)}$$

where $\beta_n$ is the ability of student n; $\delta_i$ is the overall difficulty of item i; $\tau_1$, $\tau_2$, $\tau_3$, ..., $\tau_m$ is a set of parameters associated with the transition between response categories; and m is the number of response categories provided for each item. Masters (1982, pp.163-166) used a maximum likelihood procedure to estimate the parameters in the model. Means and standard deviations of the infit (weighted) and outfit (unweighted) statistics allow a check on how well the data fit the proposed model. The advantage of the partial credit model over other models is that the parameters in the model are separable. This makes it possible to produce sufficient statistics for person ability and for item difficulty.

The QUEST Interactive Test Analysis System (Adams & Khoo 1993) which is based on the preceding equation was used to process the data. This software includes the most recent developments in Rasch measurement theory for the analysis of test and questionnaire responses as well as traditional analysis procedures. Three measures are available to check on the suitability of the tests as a measure of geometric understanding. First, the item consistency, the measure of internal consistency (p.41), used in the QUEST model is similar to Cronbach's alpha and one of the measures used to estimate the extent to which items reflect the same underlying construct. Second, the fit statistics for question estimates and student estimates are weighted and unweighted residual-based statistics and are expressions of the fit of the item to the model (p.28, p.85). They can check whether the data compared favourably with the model. Third, the infit mean square map shows the infit mean square for each item and two vertical dotted lines which enclose acceptable values. One line indicates a mean square which is 30% above the expected value and the other a mean square which is

30% below the expected value (p.23). This can determine the level of parameter fit in the model for each question. The estimates produced by QUEST are used to create a number of statistics which then allow analysis of various features of student performance and of the hierarchy of levels used to code the questions.

Mayberry's method of evaluation assesses whether a response demonstrates reasoning at the level for which the question was designed. Should a response fail to demonstrate that level of thinking it does not contribute to the assessment of the student's thinking in any way. In contrast, Gutiérrez, Jaime and Fortuny (1991) developed an assessment method in which all levels of thinking reflected in a response are evaluated. The arithmetic mean for each level is determined, and then expressed in qualitative terms, indicating complete, high, intermediate, low or no degree of acquisition. Although the Gutiérrez *et al* evaluation is a lengthier process than the Mayberry scaling method, it allows for each response to be assessed according to the van Hiele levels demonstrated therein, resulting in a possibly more inclusive picture of a student's geometrical reasoning. It also offers a way of identifying those students who are in transition between levels. Full details of the evaluation method developed by Gutiérrez *et al* are given in Chapter 7.

The next four chapters are concerned with data analysis. The first, Chapter 4, is concerned with the results of the written version of the Mayberry test. In it are presented the results by concept and for the whole test, the Guttman analysis, and a comparison between the background of the subjects and their test performance. The Rasch analysis is presented in Chapter 5, while Chapter 6 focuses intensively on the meanings or understandings students applied to the test items. This leads to an analysis of whether the Mayberry items measure the van Hiele levels for which they were designed, and whether the Mayberry success criteria are valid. Chapter 7 investigates the applicability of the Gutiérrez *et al* method of evaluation to the Mayberry test items, and how this assessment compares to the results obtained using the Mayberry evaluation method.

## Chapter 4

# RESULTS OF THE TEST*

In the previous chapter, the development of the written test and the methodological considerations incorporated in the study were described. The development, and the random distribution of the two test papers meant that each student attempted questions for four concepts. Since all the codings were to be undertaken by the writer, it was important to demonstrate consistency of the use of the Mayberry marking scheme. Consequentially, a marker who was experienced in coding student responses with the van Hiele levels agreed to act as a co-marker. The details of the approach taken are given in Appendix E. Following this, the responses of each student for each concept and each level were evaluated according to Mayberry's scaling, and tabled as four sets of quadruples. These, together with the full results for the students are detailed in Appendix F.

The purpose of this chapter is to focus on the first research theme identified in Chapter 2, namely, how does a sample of Australian students training to be primary teachers, perform on a written test version of the Mayberry questions? To assist the exploration of this theme, the chapter is divided into eight sections. The first seven sections consider the students' results for each of the individual concepts. In order, they are square, right triangle, isosceles triangle, circle, parallel lines, congruency, and similarity. In her report, Mayberry illustrated her results with details for specific items. When pertinent, details of the evaluation of responses to these items and other items of interest have been given in the results. Interview transcripts are provided when appropriate, to illustrate important features. The final section provides an overview of the results in terms of levels, as well as considering the influence of background. In the tables, results are given as numbers and (percentages).

### Square

In the Mayberry test there are sixteen scoring question parts assessing this concept. Two parts tested for Level 1, three parts for Level 2, nine parts for Level 3, while Level 4 contained two scoring parts. The items for the square were included in both Papers I and II, and, hence, all sixty-one students were tested for understanding of this concept. Table 4-1 summarises the highest level reached by each student.

---

* Aspects of this work have already been published, see Lawrie (1996), Lawrie and Pegg (1997).

Table 4-1

Highest Level Reached by Each Student for the Square (n= 61)

| No level | Level 1 | Level 2 | Level 3 | Level 4 |
|---|---|---|---|---|
| 0(0%) | 2(3%) | 51(84%) | 4(6.5%) | 4(6.5%) |

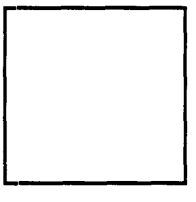
All sixty-one students achieved Mayberry's criterion for Level 1. Since the square was considered to be the most familiar concept of the seven to be tested, it was unexpected that only eight students (13%) were able to meet criteria for Level 3 or higher. Overall, three response pattern errors were identified. Each of them resulted from students correctly answering at least one Level 4 question, but failing to reach Mayberry's Level 3 criterion. There were 244 possible results for the square (i.e., sixty-one students across four levels), giving an acceptable Guttman coefficient of reproducibility for the concept of 0.99.

**Level 1**

Two questions are involved, Item 2 (identifying a figure) and Item 9(a) (discriminating figures in a group). Criterion for this level is a score of one correct response out of two. Interesting features arising in the responses to both items are detailed.

All students identified the square correctly in Item 2, thus resulting in 100% showing mastery of Level 1. The diversity in the levels of the responses is demonstrated in the following responses to Item 2:

Item 2



This figure is which of the following?

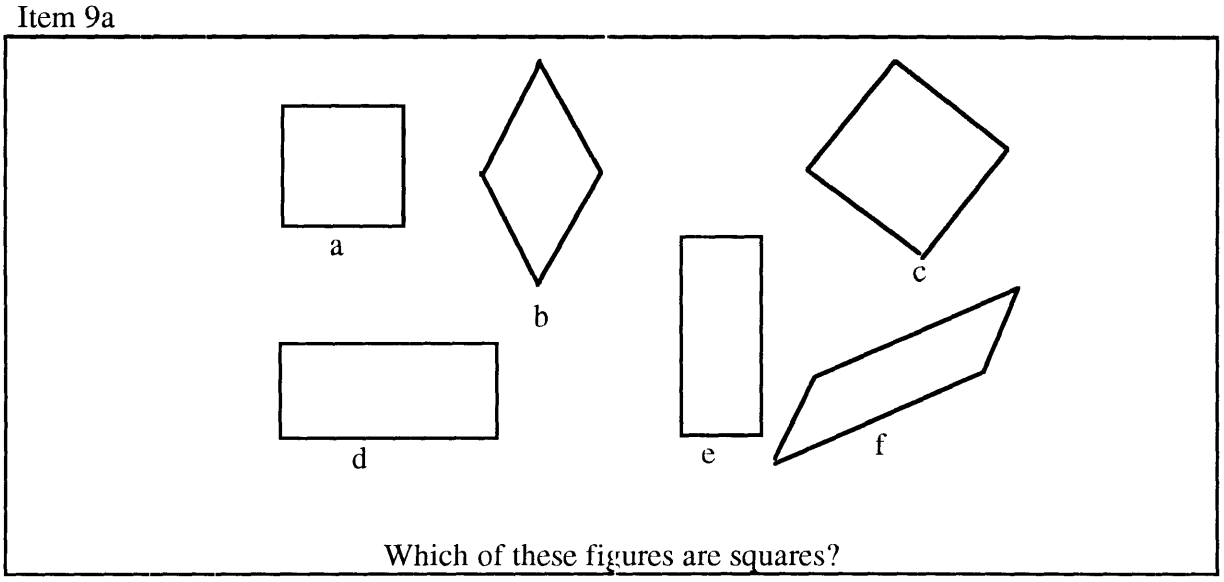
a) triangle
b) quadrilateral
c) square
d) parallelogram
e) rectangle

The majority of students (52 of the 61) gave *square* as their only response, nine students selecting other options or combinations of options (b) to (e). The difficulty of establishing the context with Level 1 questions was highlighted when three students, S14, S24, and S58, qualified their responses. A typical answer was given by S14, namely,

S14 *(b) quadrilateral, as no markings, i.e. equal length etc.*

In Item 9a, a correct response required the identification of both squares, 'a' and 'c'. All students identified correctly the standard square 'a'. However, fifteen students (25%) either failed to identify the square in a non-standard orientation 'c', or identified a non-square 'b' as a square.

Item 9a




The responses to the Level 1 questions raised the issue of the difficulty for students in establishing the context of such questions. It should be noted also, that although all students were able to recognise the square, not all students displayed the same familiarity with the square in a non-standard position.

**Level 2**

For the three question parts (criterion of 2 correct answers) testing for Level 2 knowledge, one involved awareness of properties of sides, while two tested for angle awareness. The two students (S01, S55) who did not meet criteria for Level 2 displayed knowledge of properties of sides alone, and nowhere volunteered any knowledge of angle properties. This is illustrated in the responses of S01:

Item 16

Draw a square.

a) What must be true about the sides?
b) What must be true about the angles?

S01 a) *all equal length*
b) (this was left blank)

S01 did not attempt to identify an angle in Item 23 (a), but did respond to the first part of Item 25 ( Level 3):

Item 25

| |
|---|
| A. Name some ways in which squares and rectangles are alike. |

S01 *Straight lines*

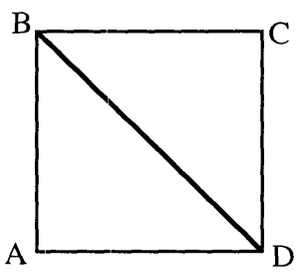
Although fifty-nine students displayed awareness of both side and angle properties for the square, the above responses indicate that weaker students are more familiar with sides than with angles. Very few students referred in their responses, to properties associated with parallel lines, diagonals and symmetry.

**Level 3**

Nine scoring question parts were used to assess whether a student is able to relate the properties of a square. Five students reached the Level 3 criterion of six correct responses. (NOTE: an additional three students reached Level 4.) The questions included the explaining of relationships/implications of properties (four questions), understanding of class inclusion (four questions), and understanding of definitions (one question).

On occasions, Mayberry designed an item to investigate knowledge of two levels, the first part asking for a listing of properties (Level 2), and the latter part requesting Level 3 reasoning. An example of this can be seen in Item 23.

Item 23




ABCD is a square, BD is a diagonal.

a) Name an angle congruent to ∠ABD

b) How do you know?

The first part of Item 23 consisted of a Level 2 task, the identification and labelling of an equal angle, while part (b) asked for a Level 3 explanation of the equality of the angles. This

question was used in the interviews to differentiate between students with Level 2 and Level 3 understanding.

Of the fifty students able to name correctly an equal angle, sixteen gave an acceptable Level 3 reason for the equality. Students assessed as working at Level 2 in the written test, did not change the level of their responses during the interviews, even after prompting. The transcript of student (S33), illustrates this.

| | |
|---|---|
| S33 (test) | a) $\angle BDC$<br>b) *not sure, can only guess* |
| S33 (int) | a) $\angle BDC$<br>b) (indicating on figure) *That* (BD) *cuts across there, so that angle* (ABD) *would equal that angle there* (BDC). |
| CL | *Is there any way you can check to make sure your answer is right?* |
| S33 | *I could measure that angle there and I could probably see that because they are 90° - being a square - they'd both be 45°.* |
| S33 | (afterthought) *I remember doing (this) in Maths. There was always a Z-shape you would look for when you were doing angles. I don't know how to describe it exactly ... that one's a parallel line and that one's a parallel line, and there was this angle* (points to alternate angles). *Do I have to explain in words?* |

This transcription not only supports the written test evaluation of a student whose responses were assessed as displaying Level 2 thinking, it also shows clearly a student attempting Level 3 reasoning, but unable to accomplish this aim This student has been given, at some time, an procedural explanation (the Z-shape) without understanding, i.e., a reduction of level has been applied.

Item 25, containing three part questions, A, B, and C, all designed by Mayberry to test for Level 3 understanding, also merits investigation.

Item 25

| |
|---|
| A. Name some ways in which squares and rectangles are alike? |
| B. Are all squares also rectangles? Why? |
| C. Are all rectangles also squares? Why? |

All students were able to list some property common to both squares and rectangles. For twenty-seven students, this was their only correct answer to Level 3 questions for the square. Sixteen students gave an acceptable response to the class inclusion part B. In this study, a student was credited with a correct response to Part C only if they had already indicated

understanding of class inclusion in Part B. For example, S03 did not gain credit for her apparently correct response to part C:

Are all squares also rectangles? Why? — *"No. All sides of squares are equal."*
Are all rectangles also squares? Why? — *"No. All sides are not equal."*

Mayberry has not offered any suggestion for her expectations of acceptable responses to any of the three parts, other than her general comment concerning Level 3 responses (1981, p.81), that credit was not given if reasons were inadequate. However, she has credited four of the nineteen students with parts (a) and (c) only.

Understanding of the meaning of a definition was tested in Item 24, in which a list of properties was given.

Item 24

| |
|---|
| Circle the smallest combination of the following which guarantees a figure to be a square. |
| a. It is a parallelogram. |
| b. It is a rectangle. |
| c. It has right angles. |
| d. Opposite sides are parallel. |
| e. Adjacent sides are equal in length. |
| f. Opposite sides are equal in length. |

Although the item was phrased so as to allow for a Level 4 response indicating understanding of sufficient as well as necessary conditions, Mayberry did not give credit for such an answer, only assessing the question at Level 3 for which it was designed. She (p.48) explained her expectations with the statement "On this level [3] a student should give definitions (since necessary and sufficient conditions are not understood, a definition may include superfluous conditions)." Since a response listing several properties can be either a Level 2 list of all known properties, or a Level 3 attempt with superfluous conditions, the students' responses to other questions were taken into account in deciding whether each answer was a Level 3 or a Level 2 response. Only two students were given credit for this question. The most common error was the omission of the necessary condition, 'a' or 'b'.

In addition to the five students achieving Mayberry's criterion for the Level 3 items, a further three students, although not reaching the Level 3 criterion, were successful at Level 4 in correctly answering one of the two items. While they registered response pattern errors between Levels 3 and 4, each failed to demonstrate understanding of a broad range of Level 3 skills. Hence, they failed to achieve Mayberry's Level 3 criterion. Their responses are discussed under Level 4.

Overall, although fifty-nine students demonstrated familiarity with the properties of a square, only five students were able to show sufficient understanding of the interaction of these properties to reach Mayberry's Level 3 criterion. While many students were able to respond correctly to the questions requiring explanations of relationships and implications, and some, the questions on class inclusion, the question requiring demonstration of understanding of the definition of the square proved to be the most difficult question.

**Level 4**

Items 45 and 46 tested whether or not a student has deductive skills. Criterion for the two items constituting this level is one correct solution. Item 45 examines a student's understanding of necessary and sufficient conditions, while the correct solution to Item 46 demands an ability to generalise in a proof While no student answered both questions correctly, four students gave acceptable solutions to one of the items. S14 was the only student to achieve Mayberry's criteria for all four levels for the square. In her responses to the Level 3 questions, she indicated understanding of a broad range of concepts. Her response to Item 45 was considered acceptable, thus meeting the criterion, but her response to Item 46 was incomplete:

Item 45

| |
|---|
| ABCD is a four sided figure. Suppose we know that opposite sides are parallel. What are the fewest facts necessary to prove that ABCD is a square? |

S14 *The sides are equal in length. The angles are 90° (or one angle is 90°)*

This response was considered to indicate sufficient understanding of necessary and sufficient conditions.

Item 46

| |
|---|
| Figure ABCD is a parallelogram, AB = BC and ∠ABC is a right angle. Is ABCD a square? Prove your answer. |

S14 *Join AC. ∴ ΔABC is isosceles (2 sides =) given.*
*∠BAC = ∠BCA (base angles isos. Δ =)*
*∠BAC = ∠BCA = 45° (180 °- 90°) (angle sum Δ)*
*∠CAD = ∠ACB alternate angles Z*
*∠ACD = ∠BAC (alternate angles)*
*∴ ABCD is a square as it has 4 angles of 90°.*

This response to Item 46, while showing skill in developing proof of right angles, failed to account for the equality of the sides.

The three students (S29, S41 and S58) who recorded response pattern errors between Levels 3 and 4, all demonstrated sound Level 3 reasoning in their answers to the Level 4 questions. Two of the students (S29, S41) were correct with Q45 and one student (S58) was correct with Q46. This knowledge of Level 3 reasoning is demonstrated in S41's written solution to Item 45, and the transcription of his interview.

S41 (test) *Line AB = Line BC. 90° angles.*

(Note: This response, in demonstrating ability to minimise one property, was taken as satisfying Mayberry's requirements for Level 4 reasoning.)

S41 (int) *Once we know that all sides are parallel ... we need to know all angles are 90° and that four sides are equal in length.*

CL (prompt) *Can you give me another answer?*

S41 *We know opposite sides are parallel and we know angles are 90°, so the figure can only be a square or a rectangle, so we need to know all sides are equal in length.*

CL *Can you do with less?*

S41 *One angle of 90°, and all sides equal.*

The transcript indicates that S41 has Level 3 reasoning, and early Level 4 reasoning also, since he is unable to minimise both side and angle properties at the same time. The interview displays the same depth of understanding as the written response.

**Summary and implications**

Many of the items testing the concept square appeared to be well constructed, and, in general, tested for the level for which they were designed. Although the concept square is studied extensively in schools in Australia, several students missed the identification of the square in a non-standard orientation. This can be taken as an indication that teachers and authors of textbooks need to be careful to familiarise their students with diagrams of figures in both standard and non-standard orientations. Additionally, some students appeared to have difficulty in determining the context of the Level 1 questions. Mayberry (1981, p. 76) observed that the Level 1 items concerning the square "were multi-purpose because responses could reflect [Level 1] or [Level 3] thought."

The greater familiarity with properties of sides than with angles demonstrated in the students' responses is supported by Pegg and Davey (1989, p.20), who commented "Many students focussed on only one property when describing plane figures - this property was invariably

the sides of the figure." While students indicated awareness of the properties of the square, most appear not to have been led into understanding of the relationships between the properties, and the symmetry of the square. The fact that only a few students answered one proof item correctly for Level 4 suggests that no student had fully mastered deductive reasoning for the concept square.

## Right Triangle

Understanding of the right triangle was tested by sixteen question parts, two for Level 1, four for Level 2, seven for Level 3, and three for Level 4. Thirty-one students attempted Paper I which contained the questions for the concept right triangle. Table 4-2 summarises the highest level reached by each student.

Table 4-2
Highest Level Reached by Each Students for the Right Triangle (n=31)

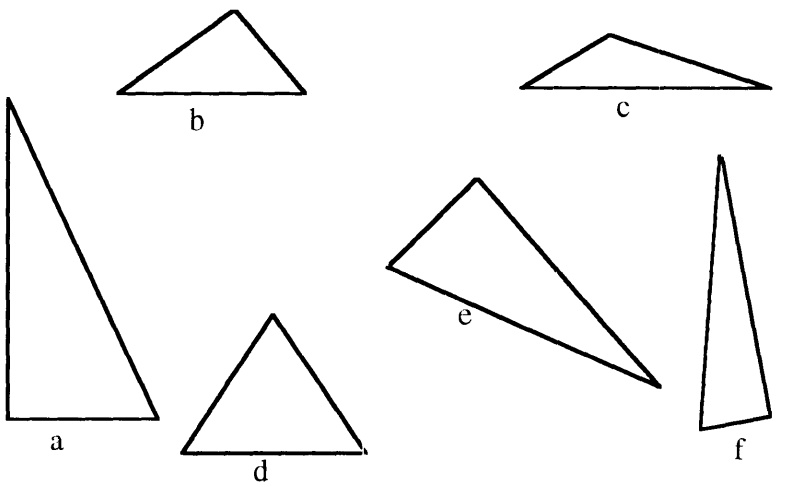
| No level | Level 1 | Level 2 | Level 3 | Level 4 |
|---|---|---|---|---|
| 1(3%) | 6(20%) | 17(54%) | 6(20%) | 1(3%) |

Thirty of the thirty one students reached Mayberry's criterion for Level 1, one student failing to reach the criterion for the first level. One response pattern error was registered when a student failed to give sufficient correct responses to the Level 2 questions, yet reached Mayberry's criterion for Level 3. One response pattern error in 124 possible results (i.e., thirty-one students across four levels) gave a Guttman coefficient of reproducibility for the concept of 0.99.

### Level 1

Two questions, Item 3 (naming a figure), and Item 10 (selecting figures from a group), tested this concept for Level 1 awareness of the right triangle. The criterion was one correct response. Whereas thirty students reached the Level 1 criterion, one student, S13, failed to show any visual awareness of the right triangle. All of this student's further attempts at questions for this concept confirmed her lack of awareness of the concept. Overall, another three students could not name the right triangle (Item 3), while nine failed to identify at least three of the four right triangles in Item 10. One identified triangle 'a' only, one 'a' and 'f', while the other seven students identified triangles 'a' and 'e'. Apart from S13, no student identified a non-right-angled triangle.

Item 10




Which of these appear to be right triangles?

The results for Item 10 again emphasise students' familiarity with a figure in the 'text-book' position.

In summary, while one student was unfamiliar with the concept, the results of the other students were mixed, a few unable to name the right triangle, while others failed to select right triangles among others. Only eighteen students were successful for both questions.

**Level 2**

This level is tested by a single question, Item 17. The item consisted of four question parts, two investigated side properties and two considered angle properties. Three question parts had to be answered correctly to achieve Level 2 criterion, and, hence, to be successful a student needed to display knowledge of both side and angle properties in a right triangle. Twenty-three students answered at least three of the four question parts correctly (there was one response pattern error).

Item 17

Does a right triangle always have a longest side?
If so, which one?

Does a right triangle always have a largest angle?
If so, which one?

For the questions testing the right triangle, errors in the responses did not show any effective difference between knowledge of side properties and knowledge of angle properties, as shown in Table 4-3.

Table 4-3
Response to Recognition of Properties of a Right Triangle (n=31)

| | |
|---|---|
| Correct response | 23(74%) |
| Sides only correct | 2(6%) |
| Angles only correct | 3(10%) |
| Missed both sides and angles | 3(10%) |

The response pattern error between Levels 2 and 3 was registered by S14, a student who, although demonstrating Level 3 and early Level 4 reasoning in her responses in both the written paper and the interview, did not appear to interpret the question as asking for the notion that the right angle must be the largest angle in a triangle Her reply to parts (c) and (d) was "*No*".

Overall, Level 2 was tested by a single item, success being measured by at least three correct responses for the four question parts. This meant students had to be able to show knowledge of both side and angle properties of the right triangle. Results do not show any difference between the familiarity of either. The failure of one capable student to respond correctly to the item raises again the question of whether there is difficulty in establishing the context of questions designed to test for knowledge at the lower levels.

**Level 3**

Seven question parts, with a criterion of five, tested whether students can demonstrate Level 3 reasoning for the right triangle. The items cover knowledge of a definition (one question), class inclusion (two questions), and relations and implications (four questions), with the more difficult questions concerning definition and class inclusion. Seven students achieved Level 3 criterion.

Responses to Item 27 in both the written test and the interviews indicate that the question parts appear to be good discriminators between students working at Level 2 and students at Level 3. This, in part, may result from the absence of a diagram, Item 27 thus requiring students to generalise in relating angle properties.

Item 27

| | |
|---|---|
| QAB is a triangle. | |
| (a) | Suppose angle Q is a right angle. Does that tell you anything about angles A and B? If so, what? |
| (b) | Suppose angle Q is less than 90°. Could the triangle be a right triangle? Why? |
| (c) | Suppose angle Q is more than 90°. Could the triangle be a right triangle? Why? |

All seven students reaching Level 3 criterion for the concept, gave three correct responses to the above item. The three students who completed Item 27 successfully, but failed to achieve the criterion for the level, were not able to define a right triangle (Item 26), nor to show understanding of class inclusion in Item 32 (a), (b), another item requiring ability to generalise. Mayberry (1981, p.83) commented on this, "In the responses to the [Level 3] questions ... students answered questions for particular figures and not for generalised ones. Thus, class relationships were not perceived."

Understanding of Level 3 relations were demonstrated clearly in the interview responses to Item 27, by students, S14 (who registered a response pattern error between Levels 2 and 3), and S41, both selected in the highest interview category. S14's responses to the three question parts are given:

S14 a) *A and B added must be 90°*
b) *Well yes, one of the other angles could be 90°.*
c) *No, because the three angles have to add to 180°, and if one is over 90°, the other two have to be less than 90°.*

In contrast, several students, including the weaker of the interviewed students (S08, S31, S33, and S59), were unable to work with a generalised triangle, and hence, resorted to quantifying a particular triangle. The transcription of S08's answer illustrates this.

S08 a) *With Q 90°, means that A and B must be 45 °.*
b) *No, it wouldn't be able to because a right triangle needs one angle to be 90°.*
c) *No, it couldn't be. I'm not sure. As a triangle is always 180°, (we were always told in school), with 90°, you can tell straight away that the other two angles are 45 ° and 45 °. No, I don't think you can have any other combination besides 45 and 45 ... because if you get a 90° angle you have to have 45 and 45.*

In summary, the Level 3 questions testing for understanding of the right triangle appear to have discriminated well between students whose awareness of properties is their highest level, and those who understand the relations and implications of the properties. That no response pattern errors occurred through students mis-understanding the context of the

questions support the notion that the items tested the level for which they were designed, and that there was good clarity.

**Level 4**

Level 4 understanding of the right triangle was tested by three items. Item 47 required a student to demonstrate ability to generalise in a proof, while Items 48 and 57 each required the construction of a proof. The criterion was a score of two correct out of three. The only student (S29) who reached the criterion, successfully completed all three items. Three students gave acceptable solutions for one of the three questions. All four students demonstrated a high degree of competency for understanding at Level 3. However, the three students completing only one proof, gave no further evidence of Level 4 reasoning for the right triangle. This suggests that a criterion of two out of three is a reliable evaluation, since it eliminates the possibility of a question previously experienced.

**Summary and implications**

In general, the questions testing for understanding of the right triangle, appeared to discriminate clearly between students reasoning at different van Hiele levels. The main problem identified was that one student recorded a response pattern error between Levels 2 and 3, and it was unclear why this was the case. For all other students in the sample, the question seemed to address the level for which they were designed.

**Isosceles triangle**

Mayberry designed nineteen question parts to assess the concept, isosceles triangle. Two question parts tested each of Levels 1 and 2, twelve question parts for Level 3, while there were three Level 4 questions. Thirty students completed the questions for the isosceles triangle (Paper II). Table 4-4 summarises the highest level reached by each student.

Table 4-4

Highest Level Reached by Each Student for the Isosceles Triangle (n=30)

| No level | Level 1 | Level 2 | Level 3 | Level 4 |
|---|---|---|---|---|
| 2(7%) | 8(27%) | 13(43%) | 6(20%) | 1(3%) |

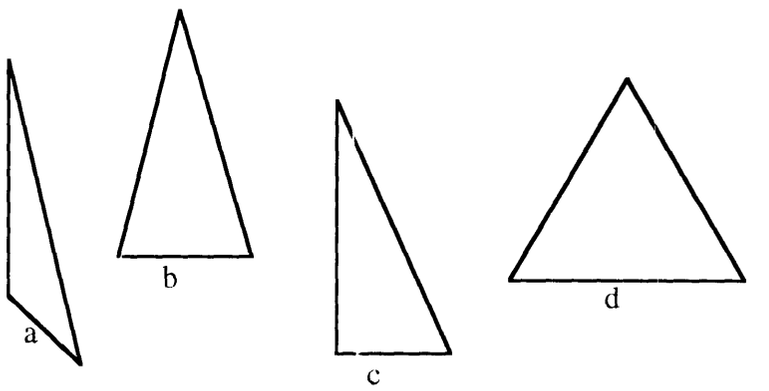
Twenty-eight students achieved Mayberry's criterion for Level 1. Two students were unable to name or identify figures of isosceles triangles. However, both students cited later that an isosceles triangle had two equal sides. In each case, the statement appears to have been recall of factual information without understanding. One student, (S57), registered a response

pattern error when she achieved the Level 3 criterion, but did not give sufficient correct responses to the Level 2 questions. This resulted in a Guttman coefficient of reproducibility for the 120 possible results (i.e., thirty students across four levels) of 0.99 for the concept.

**Level 1**

Level 1 was tested by two questions with a criterion of one correct response. All, except the two students mentioned above, correctly named the isosceles triangle (Item 4). The second question (Item 11) required the identification of both triangles 'b' and 'd' to score for this item, i.e., students being tested for Level 1 were required, for this concept, to recognise a Level 3 feature, that an equilateral triangle is also isosceles.

Item 11




Which of these figures appear to be isosceles triangles?

Fifteen students identified the correct triangles, while ten students identified only the acute-angled isosceles triangle 'b'. Some of these students rejected the notion that equilateral triangle 'd' is an isosceles triangle. For example, S42, added the qualifier, "*an isosceles Δ is one which has 2 equal sides and 2 equal base ∠s.*" This lack of understanding of class inclusion continued consistently through her responses to questions concerning the isosceles triangle. The Mayberry study does not offer any further information on the evaluation of responses to this question.

**Level 2**

The two parts of Item 18, one concerning properties of sides and one concerning angle properties, were the only test for Level 2 knowledge of the isosceles triangle. Both parts of this item had to be correct to achieve criterion, i.e., the a student needed to show knowledge of both side and angle properties to be credited with Level 2 understanding.

Item 18

| | |
|---|---|
| (a) | What can you tell me about the sides of an isosceles triangle? |
| (b) | What can you tell me about the angles of an isosceles triangle? |

Only nineteen students were successful for Level 2 (there was one response pattern error). Eight students gave a correct response for part (a), but were unable to comment satisfactorily on the angle properties, and three students were incorrect for both question parts.

S57 registered a response pattern error between Levels 2 and 3 in not giving a correct response to part (b) of this question. She appears to have mis-understood the context of the question in her answer *"They are obtuse."* Her Level 3 responses indicated understanding of the relations and implications of both side and angle properties.

Overall, the questions and criterion appeared appropriate. Student responses to Item 18 support again the notion that the properties of angles can be more difficult for students to understand than the properties of angles.

**Level 3**

This level was examined extensively, with twelve question parts testing understanding of a definition (two questions), class inclusion (five questions), and implications (five questions). Seven students achieved Mayberry's criterion of eight correct responses.

Items 28 and 29 offer an opportunity to compare the difference between two similar Level 3 items. Both questions tested whether a student can define an isosceles triangle.

Item 28

Circle the smallest combination of the following which guarantees a figure to be an isosceles triangle?

| | |
|---|---|
| a | It has two congruent angles. |
| b | It is a triangle. |
| c | It has two congruent sides. |
| d | An altitude bisects the opposite side. |
| e | The measure of the angles add up to 180°. |

This item was one of a set of questions covering four concepts (Items 24, 26, 28 and 39), which required students to demonstrate understanding of the necessary conditions for a definition, a Level 3 skill. As mentioned earlier, for the concept square, identification of too

many properties was acceptable, provided a student indicated elsewhere, Level 3 thinking. In contrast, Item 29 required the recall of a definition.

Item 29

| |
|---|
| Give a definition of an isosceles triangle. |

Students found Item 29, which allowed for recall of a memorised statement, much easier than Item 28, which required demonstration of understanding of necessary conditions for a definition. Twenty-five students gave a correct statement for Item 29, while only five were successful with Item 28.

Item 31 contained three question parts investigating whether students understood class inclusion. The first part required students to give an explanation of their answer, while the second and third parts required only a true/false response.

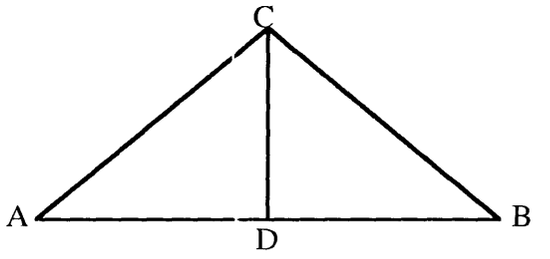
Item 31

| | |
|---|---|
| (i) | Triangle DEF has three congruent sides. It is an isosceles triangle? Why or why not? |
| (ii) | Are the following true or false? |
| | (a) All isosceles triangles are equilateral. |
| | (b) All equilateral triangles are isosceles. |

A correct response of 'false' will be given to part (ii) (a), not only by students understanding class inclusion for triangles, but also by those students who perceive isosceles and equilateral triangles as distinctly different. The final question part (ii) (b), although not requiring students to qualify their responses, should receive a correct answer of 'true' only by students understanding class inclusion for isosceles triangles. The results supported this. Five students answered part (i) correctly, twenty-five students were successful with part (ii) (a), and 7 students were correct for part (ii) (b). For nineteen of the twenty-five students responding correctly for part (ii) (a), this was their only correct response for Item 31. All except one of the nineteen students failed to reach criterion for the level.

Item 49 required students to apply their knowledge of a simple congruency proof to justify that $\Delta ABC$ is isosceles (corresponding sides AC and BC are equal). This matches van Hiele's (1957, p.239) example of Level 3 skills (implications) in the demonstration of ability to deduce the equality of angles or linear segments of specific figures, on the strength of general congruency theorems.

Item 49




ABC is a triangle. $\Delta ADC \equiv \Delta BDC$.
What kind of triangle is $\Delta ABC$? Why?

Results in this study lent support for the view that this item was a Level 3 discriminator. All thirteen students who responded successfully to this question, answered correctly several other Level 3 questions for the isosceles triangle.

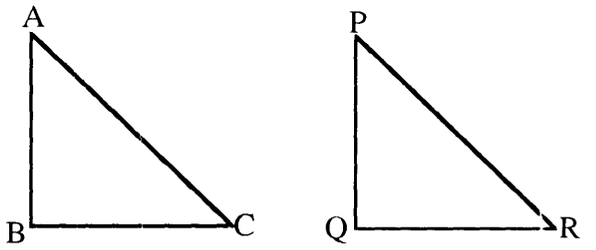
Overall, the questions appeared to result in discrimination for students' reasoning, between Levels 2 and 3, and between Levels 3 and 4. However, the results suggested that not all questions displayed the same degree of difficulty. For example, it is unclear why Mayberry included Item 29, particularly in the light of remarks concerning the validity of a similar, initially proposed item, made by van Hiele, "I can learn a definition by heart. No level. I can understand that definitions may be necessary: [Level 3]. Most pupils learn by heart." (Mayberry, 1981, p.53).

**Level 4**

Three questions, Items 48, 50, and 52, were designed by Mayberry to test whether students could work formally with isosceles triangles, i.e., could demonstrate Level 4 reasoning. Criterion for the level, as for the concept, right triangle, was two correct solutions. This was achieved by only one student, S41.

Items 48 and 52 each required the construction of a proof, the solution for Item 48 being a multi-step proof, while Item 52 required the identification of the radii of a circle as the equal sides of an isosceles triangle. To give an acceptable response to Item 50, students must have understanding of the essence of a proof. Results for the three items are discussed below.

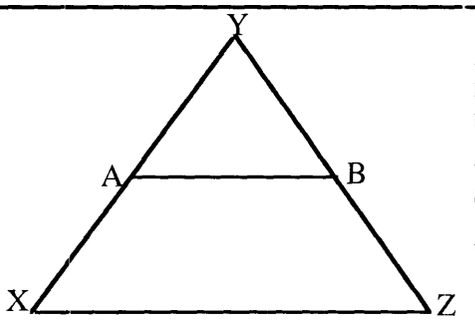
Item 48




Figures ABC and PQR are right isosceles triangles with angles B and Q being right angles.

Prove that $\angle A = \angle P$ and $\angle C = \angle R$.

Item 50




AB is the line segment with A and B the midpoints of the equal sides of the isosceles triangle XYZ.
AY = BY and $\Delta AYB$ is similar to $\Delta XYZ$ so $\angle A = \angle X$ and AB is parallel to XZ.

What have we proved?

Item 52

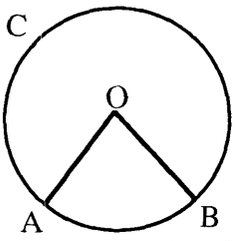



Figure C is a circle.
O is the centre.

Prove that triangle AOB is isosceles.

S41, gave acceptable solutions for all three items. He was the only student to give an acceptable response for Item 50. In his interview, a probing question for Item 50 allowed him to display the depth of his Level 4 reasoning.

S41 (written and interview)
*We have proved that in an isosceles triangle, any line segment parallel to the base will form a similar isosceles triangle.*

CL (probe)
*If A and B are not mid-points of XY and ZY, instead they divide the sides in the same ratio, what changes does this make?*

S41 (interview)
*We can move AB anywhere there and maintain similar triangles, so that say, AB is one-third of the way along, we still have similar triangles. If A was one-third and B was half, AB is no longer parallel, so the triangles are not similar.*

The results for the items indicate that the students did not find the three questions of similar difficulty, Item 52, the two-step proof question, having the greatest rate of success, and Item 50, the least. Table 4-5 compares the successful response rates for the three Level 4 items.

Table 4-5
Number of Students Successful with Level 4 Items (isosceles triangle)

| | |
|---|---|
| Item 48 correct | 3 |
| Item 50 correct | 1 |
| Item 52 correct | 10 |
| One item correct | 11 |
| Two or more items correct | 1 |

Of the ten students giving an acceptable response for Item 52, only three achieved criteria for Level 3. These results support the finding above for right triangles, that two out of three correct responses is a discriminating criterion.

**Summary and implications**

Many of the questions testing for understanding of the concept isosceles triangle, appear to discriminate clearly between reasoning at the various van Hiele levels. However, assessment of the responses has given rise to several issues. First. students finding questions about the properties of sides easier than questions about the properties of angles supports Davey and Pegg's (1991, p.1) comment that the angle concept is particularly difficult in that it requires the student to integrate a number of different facets, and Mitchelmore's (1993, p.405) finding that children have several problems in abstracting the angle concept. Second, a criterion of 100% discriminates against students mis-interpreting a question. Third, the stating of a recalled definition does not necessarily test for understanding of a concept. Pegg and Davey (1989, p.23) recorded that, in describing a figure as best they could, many of the students "gave correct mathematical definitions as their descriptions of the figures. ... Unfortunately, the significance of their answers was beyond their comprehension and did not reflect some special insight into the figures concerned." Finally, three questions with a criterion of two correct responses, as shown in the Level 4 questions for this concept, appears to give a more reliable evaluation of reasoning than two questions with a criterion of one. These issues are discussed in Chapter 6.

## Circle

There were twenty question parts designed to test for understanding of the circle concept. As with the other concepts, two questions were used to determine whether a student could recognise a circle (Level 1). Six questions parts tested for Level 2 understanding, ten parts tested for Level 3 and two questions tested for Level 4. Thirty-one students attempted the twenty questions in Paper I. Table 4-6 summarises the highest level reached by each student.

Table 4-6

Highest Level Reached by Each Student for the Circle (n=31)

| No level | Level 1 | Level 2 | Level 3 | Level 4 |
|---|---|---|---|---|
| 0(0%) | 4(13%) | 6(19%) | 16(52%) | 5(16%) |

More students achieved criteria for Levels 3 and 4 for the circle, than for any other concept. This suggests that the students have a greater degree of understanding of the concept circle than of the other six concepts. Three students registered response pattern errors, each student achieving Level 3 criterion, but failing to demonstrate sufficient knowledge of properties (Level 2). This gave a Guttman coefficient of reproducibility for the 124 possible results (i.e., thirty-one students across four levels), of 0.98.

### Level 1

This level was tested with two questions, the criterion being one correct response. All students successfully named the circle (Item 5), and hence all students registered success at Level 1. Item 12 required students to identify two circles from four figures. Three failed to identify the smaller of the two circles.

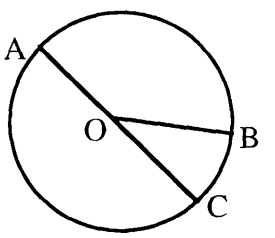
### Level 2

Items 19 and 33 assessed whether a student is working with Level 2 skills. Together, the two questions contain six scoring parts, the criterion being five correct responses. All except seven students reached the criterion for the level. However, of these, only four were unable to demonstrate skills higher than Level 1. The other three students reached Level 3 criterion, and, hence, registered response pattern errors.

Item 19 required the identification of three parts of a circle, two radii and a diameter, two of which had to be named correctly to score one credit point. By comparison, Item 33 required students to state why each of five figures is or was not a circle. Each figure scored one credit

point, giving a total of five points for the question. Mayberry did not offer any explanation why Item 33 should be worth five times as much as Item 19.

Item 19

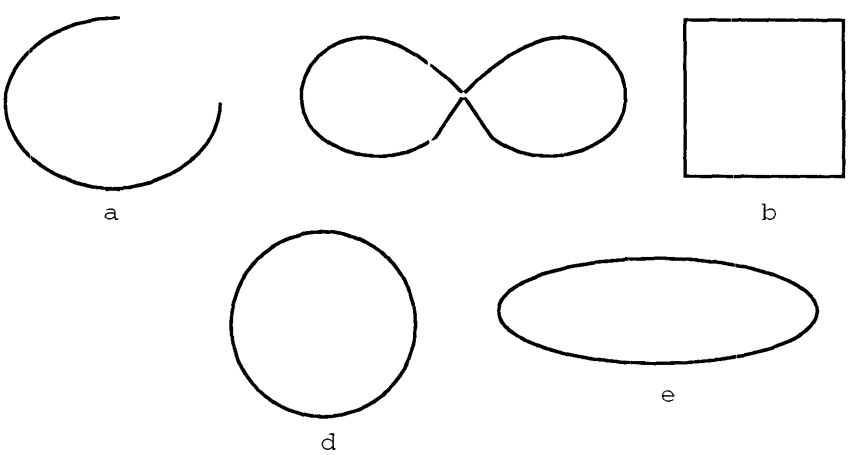



This figure is a circle. O is the centre.

Name the following line segments.
OB is a _____ of the circle.
OC is a _____ of the circle.
AC is a _____ of the circle.

Eight students failed to give at least two acceptable responses for Item 19. Four of the students showed little awareness of circle properties, and, hence, did not reach Level 2 criterion. The other four students, although demonstrating knowledge of these properties in later responses, appeared not to comprehend the intention of the item, often not giving any answer. As a consequence, two of the students (S22 and S30) recorded response pattern errors. The third response pattern error occurring between these levels resulted from S02 not completing Item 33. S02 appeared not to have understood the intention of the question parts.

Item 33




(a) to (e) Tell why each of the figures is or is not a circle.

For Item 33, while most students listed all known properties in their responses, a few students answered questions (a) to (e) with Level 1 (visual) answers, failing to indicate any knowledge of circle properties. Such answers included: *"does not meet", "not a full shape", "looks like a flat tyre", "crosses in the centre", "it's a square", "looks like one* (circle)", and *"it is oval"*.

The failure of three students to achieve Mayberry's criterion for this level while successfully demonstrating Level 3 reasoning, and, hence, registering three response pattern errors, again raises the question of the difficulty in establishing the context of Level 2 items.

**Level 3**

The thinking displayed in responses to ten question parts, with a criterion of six, determined whether a student was credited with Level 3 reasoning for the concept circle. The questions were classified by Mayberry as requiring demonstration of understanding of a definition (one part), of relations (five parts), and of class inclusion (four parts). Twenty-one students were successful in reaching the criterion, i.e., only ten students were unable to respond satisfactorily to Mayberry's questions testing for Level 3 thinking for the circle.

The understanding of a definition was tested by another of Mayberry's multi-level questions. In Item 33 (above), the first five parts could be answered at either Level 2 or Level 3, and required to be evaluated for both levels. In her interviews, Mayberry was able to prompt for a Level 3 response. This is not possible in a written test. Hence, a final part (f) was added to the original question, giving students the opportunity to demonstrate their Level 3 ability to generalise, should they have failed to do so in parts (a) to (e).

Item 33 (text only)

| | |
|---|---|
| (a) to (e) | Tell why each of the figures is or is not a circle. |
| (f) | Can you give a general rule to fit all the above answers? |

Of the sixteen students given credit for a Level 3 response, only five did so in their responses (a) to (e), i.e., the responses (a) to (e) were most commonly at Level 2, justifying the inclusion of part (f).

Many students gave acceptable answers to the question parts in Items 34 and 35. Item 34 containing four scoring parts, tested for understanding of class inclusion, while Item 35, also with four scoring parts, tested for knowledge of the relations between properties.

Item 34

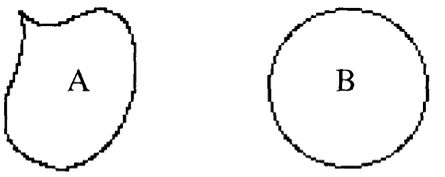


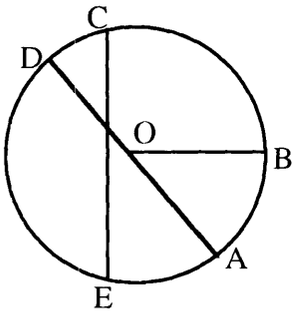

Figure A is a simple closed curve.
Figure B is a circle.

(a) Is figure B a simple closed curve?
(b) How are these figures alike? How are they different?
(c) (T-F) All simple closed curves are circles.
(d) (T-F) All circles are simple closed curves.

Most of the twenty-one students reaching the Level 3 criterion scored well for this item, eighteen giving four correct responses, and one giving three. However, parts (a), (c) and (d), being yes/no or true/false questions, have a 50% chance of a correct guess. The possibility of students making correct guesses without the necessary knowledge is illustrated in the attempts of three of the four students who were successful at Level 1 only. Two scored two correct responses, and one student scored three correct responses for the question parts in Item 34. The remaining student did not attempt the question.

As with Item 34, most students did not find Item 35 difficult, twenty-three students giving at least three correct responses. Three students failed to give reasons for their answers, and, hence, failed to score for the item. Among these was S31, one of the interviewed students, who, in general, showed good Level 2 and early Level 3 reasoning for all concepts. S31 reached Mayberry's Level 2 criterion for the circle, but not the Level 3 criterion. The transcript of his interview for parts 3 and 4 of Item 35, in showing his reasoning, illustrates one of the problems associated with a written test, i.e., the difficulty in having the students write down all that they know.

Item 35




This figure is a circle with centre O. Would the following be: a) certain, b) possible, c) impossible. Give reasons for your answer.

1) OB=OA 2) OD=OA

3) 2OB=AD 4) AD=EC

S31 (Part 3) (A line extending BO to the circumference was drawn first)
*Certain, they are both diameters of a circle. The diameter of a circle goes through the centre O to the other side in a straight line.*

S31 (Part 4) *Impossible because CE doesn't go through the centre O, so therefore, CE is going to be shorter than AD, which is longer because it goes through O and it goes to both sides.*

CL *Is there any way you can check your answer?*

S31 *Well, measurement by ruler.*

CL *Suppose I tell you that AD is the same length as EC, what are your comments?*

S31 *I'd say it couldn't be. I can probably see with my eyes it's not, because I know CE doesn't go through the centre O. Can I use a ruler?*

The transcript supports the assessment of S31, showing him as a student who, in knowing the properties of radii, diameters and chords, is attempting early relational skills. It is interesting that when S31 was confronted with a conflicting situation, he reverted to the more comfortable Level 2 reasoning of proof by measurement.

Overall, most of the students were able to answer the Level 3 items correctly, possibly indicating familiarity with concept. However, the design of some of the questions could allow students to gain credit from a 'guess'.

**Level 4**

Items 51 and 52 together test for Level 4 skills in the concept circle. Students are assessed as having formal skills if they can correctly answer one of the two items. Item 51 requires demonstration of understanding of the essence of a proof, while Item 52, also used by Mayberry to test for knowledge of the isosceles triangle, and discussed already in that section, requires the development of a two-step proof. Five students reached the Level 4 criterion, each scoring one correct answer. One student answered Item 51 correctly, and four gave correct responses to Item 52.

A satisfactory response to Item 51, for example, 'equal angles are subtended at the centre of a circle by equal arcs', cannot be given without a student being able to overview the whole proof, i.e., being able to reason formally. S19, demonstrated this ability in his response, *"that the chords are equal due to having the same angle at the radii and centre."* S19 failed to complete the proof for Item 52.

Item 51

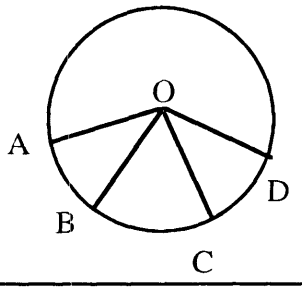



Figure O is a circle, O is the centre.
$\angle AOB = \angle COD$.
Since AO, BO, CO and DO are radii,
$\Delta$ AOB $\equiv$ $\Delta$COD, so AB = CD.

What have we proved?

Of the four students successful for Item 52, one did not attempt Item 51, while three selected only part of the proof, indicating Level 3 reasoning. The response, 'that $\Delta$ABO and $\Delta$CDO are congruent', i.e., a selection of one of the statements given in the proof, indicated that, although students were able to relate properties, they were unable to comprehend the totality of a proof.

The above comparison of the responses for the two items suggests, as with the items testing Level 4 for isosceles triangle, that there may be a difference in the degree of difficulty between the items.

**Summary and implications**

The concept circle appears to be better understood than the other six concepts tested, with 68% of the students being able to achieve Mayberry's criterion for Level 3 or better. The transcriptions of the interviews indicate that the levels of reasoning determined for the written test, are stable. However, the responses to Item 33 show clearly that interview questions can need modification if they are to elicit the required depth of reasoning in a written test. There is no discernible reason for the three response pattern errors occurring between Levels 2 and 3. The success of weaker students with true/false questions, e.g., parts (c) and (d) of Item 34, again raises the issue of whether such questions are of value in testing for Level 3 understanding.

## Parallel Lines

Fifteen question parts tested for understanding of the concept parallel lines, two each for Levels 1 and 2, nine for Level 3, and two for Level 4. Thirty students completed the questions (Paper II) for this concept. Table 4-7 summarises the highest level reached by each student.

Table 4-7

Highest Level Reached by Each Student for the Concept Parallel Lines (n=30)

| No level | Level 1 | Level 2 | Level 3 | Level 4 |
|---|---|---|---|---|
| 0(0%) | 5(17%) | 24(80%) | 0(0%) | 1(3%) |

The results for parallel lines indicate that, of the seven concepts, this one was the least understood, with only one student demonstrating understanding greater than Level 2. Mayberry (1981, p.86) commented that her students found this concept to be one of the more difficult. The student achieving Level 4 criterion, S53, was unsuccessful at Level 3, and hence, registered a response pattern error. This gave a Guttman coefficient of reliability of 0.99 for the 120 possible results (i.e., thirty students across four levels) for the concept parallel lines.

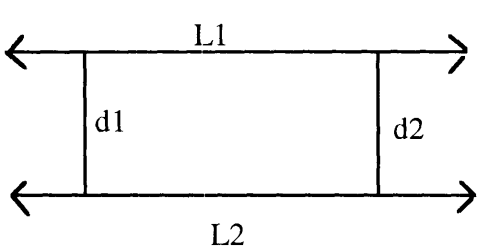
**Level 1**

Two questions with a criterion of one correct response, tested for visual recognition of parallel lines. All students responded correctly to at least one of the questions. Twenty-eight students successfully named the parallel lines in Item 6. Item 13 required the identification of two sets of parallel lines, one pair being in the horizontal 'textbook' position. Only two students failed to identify the non-horizontal pair of parallel lines. Both these students reached the criterion for Level 2. For the two students who failed to name the parallel lines (Item 6), the identification of the two sets of parallel lines in Item 13 was their only correct answer for the concept.

**Level 2**

Item 20 was the only question set by Mayberry to test Level 2 skills for parallel lines. It contained two scoring parts, both of which had to be answered correctly to reach the criterion. This item tested for awareness that parallel lines are everywhere the same distance apart, and that they never meet.

Item 20




If d1 is equal to d2, what is true about L1 and L2?

If d1 is not equal to d2, what is true about L1 and L2?

Five students failed to reach the criterion. Of these, the two who attempted the questions, both used the incorrect phrase *"they are equal."* All five students made little or no attempt with the Level 3 questions, demonstrating a lack of familiarity with the concept.

**Level 3**

Understanding of Level 3 skills, i.e., understanding of a definition (two parts), relationships (five parts) and implications (two parts), was tested by nine question parts. For this concept, Level 3 questions were not well attempted. No student recorded more than three correct answers, well short of the criterion of six. In particular, students appeared to be unaware of the notion of skew lines, and, hence, that if non-intersecting lines are to be parallel, they must be in the same plane. This feature, a condition of parallelism included by Mayberry for Level 3, is implicit to obtaining the correct responses to Items 38 and 39, questions which test for understanding of the relevant relationships and for sufficient understanding in the definition of parallel lines

Item 38

| |
|---|
| Are these lines or line segments parallel, a) always b) sometimes c) never?<br>Give reasons for your answers.<br><br>A Two lines which do not intersect.<br>B Two lines which are perpendicular to the same line.<br>C Two line segments in a square.<br>D Two line segments in a triangle.<br>E Two line segments which do not intersect. |

No student included the possibility of skew lines in their responses to parts A, B or E. The three students giving acceptable responses for part E, all based their answers on the possibility of the two line segments meeting if extended indefinitely.

Item 39 was included in the interviews, in part, to investigate this apparent lack of awareness that parallel lines are co-planar. No students included condition (c), 'they are in the same plane', in their responses for the written paper. Interview responses for Item 39 illustrate the extent of this lack of understanding.

Item 39

| |
|---|
| Circle the smallest combination of the following which guarantee that two lines are parallel.<br><br>a) They are everywhere the same distance apart.<br>b) They have no points in common.<br>c) They are in the same plane.<br>d) They never meet. |

In the interviews, the students indicated initially the answer to be one or more of the selections, (a), (b), and (d). All six students omitted (c), some appearing to ignore the option, others expressing uncertainty as to the meaning of the term 'plane'. When this occurred, the term was explained. Each student was then prompted with the notion of skew lines, examples being pointed out in the interview room. The prompting caused only one student, S14, whose initial answer was (a), to change her response. S14, having completed written test Paper I, had not attempted Item 39 before the interview. Her interview reveals Level 3 reasoning in that she recognised immediately the significance of the condition, that parallel lines must be coplanar.

CL *Suppose you consider the two lines in this room* (a pair of skew lines, ceiling and floor, are indicated). *What can you tell me about them?*

S14 *Well, they're both straight, but they're not parallel. They have to be going in the same direction,* (student pauses) *in the same plane.*

CL *Do you now wish to change your original answer?*

S14 *Yes. The answer should be (a) and (c).*

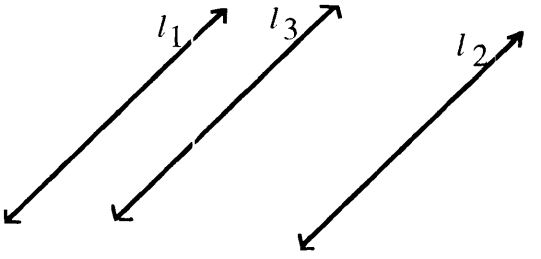
S31 and S41, in their interviews, recognised that the skew lines were not 'in the same plane', yet would not change their original answers to include the condition. The other three students found the notion of the skew lines interesting, but could not 'take on board' the concept of planes.

For parallel lines, the students' responses indicated that they were not familiar with the concept at Level 3, and, in particular, that they had not considered the 3-dimensional notion that, to be parallel, lines must be co-planar. Mayberry (1981, p.83) found a similar degree of unawareness of the characteristic with her students, commenting "thirteen of the fifteen students who missed this question (Item 39) about parallel lines omitted 'in the same plane'."

**Level 4**

Items 53 and 54 tested whether or not a student has deductive skills for the concept parallel lines. Criterion for the level was one correct solution. Item 53 required understanding of the essence of a proof, while Item 54 asked students to construct a multi-step proof. No student gave an acceptable answer for Item 53. One student, S53, was credited with an acceptable response to Item 54, thereby registering a response pattern error between Levels 3 and 4. He scored only two correct responses for the Level 3 questions.

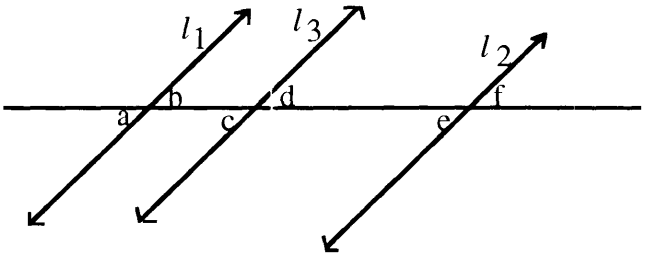
Item 54




Prove: If $l_1$ is parallel to $l_2$ and $l_2$ is parallel to $l_3$ then $l_1$ is parallel to $l_3$ .

In S53's solution for Item 54, he constructed a transversal (as shown below), identified angles with pronumerals, and then indicated the equality of the vertically opposite angles at the commencement of his proof.

Item 54



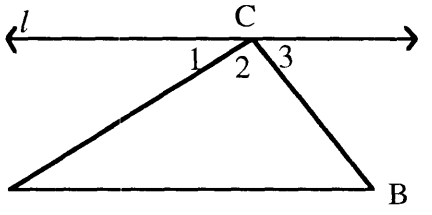

$\angle a = \angle b,\ \angle c = \angle d,\ \angle e = \angle f$ (X) [This symbol is used by the student as shorthand for vertically opposite angles]

His proof was:

S53 $\angle a = \angle b,\ \angle c = \angle d,\ \angle e = \angle f$ (X diagram, indicating vertically opposite angles)
*If* $l_1 \parallel l_2$ *and* $l_2 \parallel l_3$ *then* $\angle b = \angle c$ *and* $\angle d = \angle e$ (Z) [This symbol is used by the student as shorthand for equal alternate angles]
$\therefore\ \angle b = \angle d$ *and* $\angle a = \angle c$, *making* $l_1 \parallel l_3$

While this is not an elegant proof, the solution was considered to show the essence of the construction of a proof (early Level 4), and, hence, was deemed to meet Mayberry's requirements. Overall, S53 did not show strength in his responses to Level 3 questions, reaching criterion for the concept isosceles triangle only. In this, he demonstrated sound mastery of the understanding of the concept at that level. However, his responses for the other concepts showed good Level 2 knowledge, but only early Level 3 understanding. This is illustrated in his answer to Item 53.

Item 53




Line $l$ is parallel to AB.
Because of properties of parallel lines we can prove that $\angle 1 = \angle A$ and $\angle 3 = \angle B$.
Now, $l$ is a straight angle (180°).

What have we proved?

S53 $\angle 2=90° + \angle 1=45° \ \& \ \angle 3=45°$
*ΔABC is a right triangle*

His need to quantify the angles in Item 53 is considered to indicate Level 2 reasoning. This contrasts with his acceptable attempt with Item 54 which provides a Level 4 answer. It is difficult to reconcile this pattern of responses.

**Summary and implications**

The results for the concept parallel lines indicates it to be the most difficult of the seven concepts tested. Only one student was able to be successful beyond Level 2. While almost all students identified parallel lines in both standard and non-standard positions, the majority of students showed little or no understanding of the intention of the Levels 3 and 4 questions. In particular, the students appeared to be unfamiliar with the notion that parallel lines need to be co planar, and unaware of the existence of skew lines.

## Congruency

Eighteen scoring question parts were designed by Mayberry to test for understanding of the concept congruency. Two question parts tested for Level 1, four parts for Level 2, ten parts for Level 3, and two parts tested for Level 4 understanding. The congruency questions occurred in written Paper I which was attempted by thirty-one students. Table 4-8 summarises the highest level reached by each student.

Table 4-8
Highest Level Reached by Each Student for the Concept Congruency (n=31)

| No level | Level 1 | Level 2 | Level 3 | Level 4 |
|---|---|---|---|---|
| 0(0%) | 10(32%) | 11(35.5%) | 1(3%) | 9(29%) |

More students achieved criterion for Level 4 for congruency than for any other concept. The occurrence of eight response pattern errors was also greater for congruency than for the other concepts. One error occurred for a student who was able to reach Mayberry's Level 3

criterion, yet gave insufficient correct responses to the Level 2 questions. For the remaining seven response pattern errors, the students were able to meet Mayberry's criterion of one correct answer for Level 4, but were unable to meet the Level 3 criterion. This number of response pattern errors, namely, 8 in 124 possible results (i.e., thirty-one students across four levels), gave a Guttman coefficient of reliability, 0.94, a value still above Torgenson's (1967) standard value of 0.90, but the lowest of the seven concepts.

As in the preliminary study, there were no indications in the responses, of students failing to understand or mis-using the term 'congruency'. Mayberry (1981, p.80) explained the term congruent to her students when she found "many students were informally (and incorrectly) using the word *equal*."

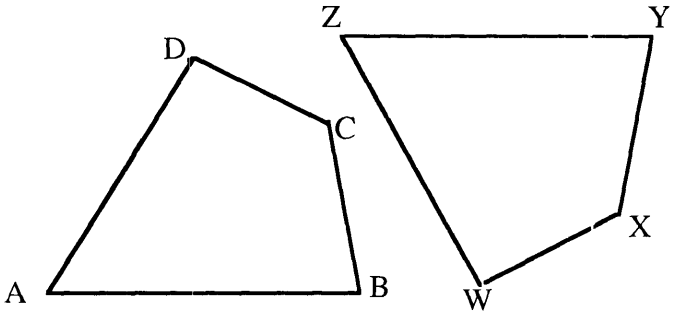
**Level 1**

Mayberry's criterion of one correct response for the two questions testing visual awareness of congruency was reached successfully by all students. Item 7, answered correctly by twenty-six students, tested for ability to identify congruency, and Item 15 (twenty-eight correct responses) required students to select a second congruent quadrilateral from a group of four figures. The wording of Item 15 caused difficulties in the preliminary study, and, hence, had been altered to read 'Which figure appears to be congruent to A?' Of the three students failing to give a correct response to Item 15, one student did not answer any questions on congruency after Item 7. The remaining two students, who gave incorrect responses, both reached the criterion for Level 2. There is insufficient information to determine whether the new format of the Item 15 is now satisfactory.

**Level 2**

Item 22 which contains four question parts, is the sole item set to test for Level 2 in the concept congruency. As with the concept right triangle, two question parts related to the properties of sides and two to the properties of angles. A criterion of three correct responses requires that at least one property for each, sides and angles, must be identified. Eleven students (there was one response pattern error) failed to achieve Mayberry's criterion (there was one response pattern error). Of these, seven scored zero or one, and four gave two correct answers.

Item 22




These are congruent figures.

(i) What is true about their sides?
(ii) AD = _____

(iii) What is true about their angles?
(iv) ∠B = _____

Responses to this item again support the notion that the students are more familiar with the properties associated with sides than with those associated with angles. Almost two-thirds of the incorrect responses were for the latter two question parts, i.e., those asking about the properties of angles. A few students appeared to misunderstand the intention of the first and third parts of the question. This may be the reason for the response pattern error registered by S06. Overall, S06 demonstrated confident Level 3 reasoning in all concepts, including congruency. Her responses to Item 22 were:

| | | | |
|---|---|---|---|
| What is true about their sides? | S06 | *they are unequal* | $AD = WZ$ |
| What is true about their angles? | S06 | *they add to 360°* | $\angle B = \angle Y$ |

The responses to the Level 2 question parts indicate that students find it easier to answer parts (ii) and (iv), i.e., where they were asked for a specific result rather than a descriptive phrase.

**Level 3**

Ten question parts test for understanding of relations between properties (four parts), their implications (five parts), and class inclusion (one question part). Only three students achieved Mayberry's criterion of six correct responses for the level. Another seven students, in producing proofs for the Level 4 questions (and, hence, recording response pattern errors), demonstrated understanding of relationships between properties, and the implications.

The question parts are all within a group of multi-purpose questions, Items 41 to 44. While the items test for Level 3 knowledge across several concepts, they focus particularly on understanding of congruency and similarity. No diagrams are given. There is, however, a difference in the phrasing of the questions. Items 41 and 43 specify triangles ABC and DEF, whereas, Items 42 and 44 refer to general figures. Many students drew figures to facilitate their efforts in formulating a response. Mayberry designed Items 43 and 44 to test specifically for knowledge of congruency.

Item 43

| | | | |
|---|---|---|---|
| $\Delta$ABC is congruent to $\Delta$DEF (in that order). | | | |
| Are the following | a) certain | b) possible, or | c) impossible? |
| Give reasons for your answers. | | | |
| a) | AB = DE | c) | $\angle A < \angle D$ |
| b) | $\angle A = \angle E$ | d) | AB = EF |

For Item 43, nineteen students failed to offer any reasons. Of the twelve who did, seven gave acceptable responses for at least half of the question parts. In contrast, twenty-three students failed to offer explanations for Item 44. Only three of the eight students attempting reasoning, gave acceptable responses to at least half of the question parts.

Item 44

| | | | |
|---|---|---|---|
| Will figures A and B be congruent | I-always | II-sometimes, or | III-never? |
| Give reasons for your answers. | | | |
| A | | B | |
| a) a square | | a) a triangle | |
| b) a square with a 10cm side | | b) a square with a 10cm side | |
| c) a right triangle with a 10cm hypotenuse | | c) a right triangle with a 10cm hypotenuse | |
| d) a circle with 10cm chord | | d) a circle with 10cm chord | |
| e) a $\Delta$ similar to B | | e) a $\Delta$ similar to A | |

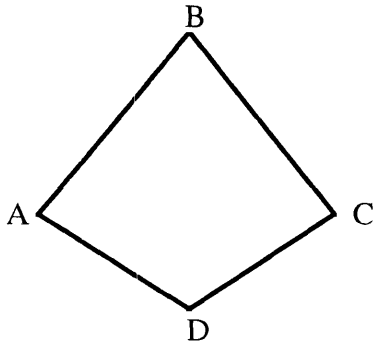
The difference between the success rates for the two types of questions match the difference in the phrasing of the two items, more students giving correct responses to the more direct Item 43 than to Item 44.

Overall, the question parts for Items 43 and 44 were not answered well. It is not clear whether the poor results are a reflection of the question types and/or their suitability to the format of a written test, or a reflection of the concept being tested.

**Level 4**

Two questions, Items 55 and 56, were designed by Mayberry to determine whether students can reason at Level 4, for the concept congruency. Both questions test whether students can construct a proof. The criterion for this level is one correct solution. Nine students achieved the criterion, four giving acceptable solutions for both questions, while five were correct for Item 56 only. Seven of the nine students failed to demonstrate success at Level 3, and, hence, registered response pattern errors between Levels 3 and 4.

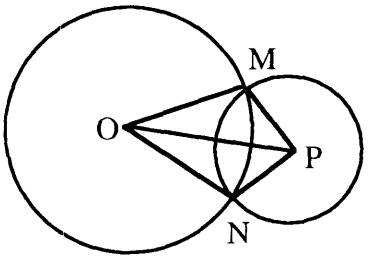
Item 55




In this figure AB and CB are the same length. AD and CD are the same length.

Will $\angle A$ and $\angle C$ be the same size? Why or why not?

Item 56




These circles with centres O and P intersect at M and N.

Prove: $\Delta OMP \equiv \Delta ONP$.

Item 55 can be solved by a number of different techniques. One solution to the problem involves the construction of a pair of congruent triangles by joining BD. The proof of congruency of these triangles then becomes an instrument used within the solution of the problem. By contrast, the diagram for Item 56 clearly delineates triangles OMP and ONP. The solution requires solely, the identification of three equal pairs of corresponding sides, thus showing congruency of the triangles.

The interviews of two students, S33 and S14, indicate Item 55 can be used to discriminate clearly between students at different levels. S33 is a student demonstrating confidence at Level 2, together with some early Level 3 skills, while S14 is a capable student, frequently showing Level 4 ability.

S33 *Yes, they look as if they are equal. I could measure them.*

CL *Can you prove this in some other way?*

S33 (She joined AC and BD, marking the point of intersection as O)
*AC and BD make a right angle. AO is equal to OC.*

*CL* *Can you tell me why?*

S33 *It has to be so.*
(S33 then drew a diagram in which the diagonals were not perpendicular, i.e., the diagram was no longer a kite)
*It won't work, the diagram, if BD and AC don't make right angles.*

With further probing, S33 became confused, and reverted to visual statements about the angles looking the same. In contrast, the transcription of S14's interview shows her confidence in solving the problem. She had no difficulty in producing two different proofs when prompted.

S14 *Well, I'll join a line from B to D. It looks easier. It's easier to handle two shapes than a whole huge shape that you're not used to looking at . . . so the triangles are congruent* (marks three pairs of corresponding sides equal), *therefore the angles have to be the same because those two angles* (indicates A and C) *are constantly equal.*

CL *Can you prove this in another way?*

S14 *Join AC instead of BD. We've got two isosceles triangles. Let angle BAC equal 'x', so angle BCA is 'x' - you know they're going to be equal because they are the base angles of an isosceles triangle, and the base angles are equal. The same with the other side* (marks angles CAD and ACD equal to '$y$'). *So angles A and C are both '$x + y$'.*

S14 was unable to produce valid reasoning when presented with a conflicting situation, resulting in an evaluation of her level of understanding as early Level 4. The transcript of S33 indicates she has confidence working with properties (Level 2), and is trying unsuccessfully, to find relationships between the properties.

In summary, the results suggest there is a difference between the degrees of difficulty of the two questions designed by Mayberry to test for Level 4 skills in writing a proof. However, this does not account for all seven response pattern errors. While three of the students, unable to give sufficient correct responses to the Level 3 items, were able to answer Item 52 only, the other four students gave correct solutions to both Items 51 and 52.

**Summary and implications**

The analysis of the responses to the questions testing congruence reinforce some of the issues already noted for earlier concepts. Again, a response pattern error occurred between Levels 2 and 3. In this case, a student achieved the Level 3 criterion but misinterpreted the intention of some of the Level 2 question parts. Additionally, the results supported the feature noted in concepts already analysed, that the concept of angles is more difficult than the concept of sides of figures.

There are two possible indicators of causes for the occurrence of the large number (seven) of response pattern errors between Levels 3 and 4. First, Items 41 to 44 with their requirement for working with generalised figures, appear to give more difficulty than other Level 3 questions. Second, there appears to be a difference between the degrees of difficulty for the two questions testing for Level 4 skills. Whereas five students gave acceptable solutions for Item 55, nine were successful for Item 56, suggesting that the spontaneous recognition of the need to construct triangles before undertaking congruency requires a deeper overview of the power of congruency.

**Similarity**

Mayberry designed twenty question parts to test for understanding of the concept similarity. Two question parts were designed to test for Level 1, four for Level 2, thirteen for Level 3, and one for Level 4. The development of a replacement Level 4 item for the written test, did not alter Mayberry's balance of questions. The concept similarity was tested in the second written paper, and, hence, thirty students attempted the questions. Table 4-9 summarises the highest level reached by each student.

Table 4-9

Highest Level Reached by Each Student for the Concept Similarity (n=30)

| No level | Level 1 | Level 2 | Level 3 | Level 4 |
|---|---|---|---|---|
| 0(0%) | 13(43%) | 12(40%) | 3(10%) | 2(7%) |

Although all students demonstrated an awareness of similarity in achieving the Level 1 criterion, less students were able to proceed to Levels 2, 3 and 4 than in any other concept. Two response pattern errors were registered, one student meeting Level 3 criterion while unable to give correct responses to the Level 2 items, and one student giving an acceptable response to the Level 4 item, but unable to achieve Mayberry's Level 3 criterion. For the 120 possible results (i.e., thirty students across four levels), this gave a Guttman coefficient of reproducibility of 0.98.

**Level 1**

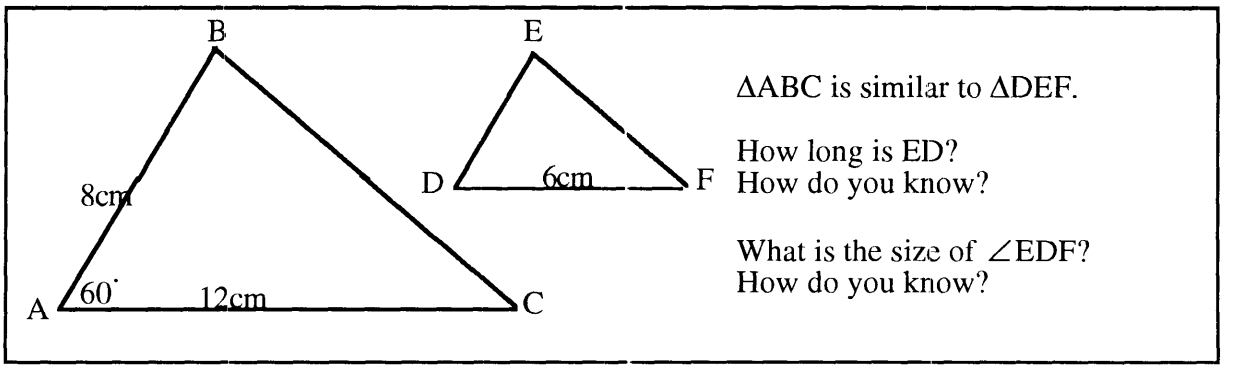
Two questions test for recognition of similarity, the criterion of one correct response being the same as for Level 1 with the other concepts. Item 8 required students to recognise and name the concept similarity, and Item 14 required students to identify the pair of similar figures in a group of four. All students reached the criterion for the level, one student failing

in Item 8, and four in Item 14, indicating that the notion of similarity is familiar to students from a wide range of backgrounds (detailed in Chapter 3).

**Level 2**

Level 2 awareness of properties for similarity is tested by four question parts, two asking about properties of sides of figures and two about angle properties. Criterion for Item 21 is again three correct responses, achievement requiring demonstration of knowledge of both sets of properties. Sixteen students achieved Mayberry's criterion for Level 2 (there was one response pattern error).

Item 21




Of the fourteen students who did not meet Level 2 requirements, four were incorrect for all questions parts, one gave correct numerical answers but no reasons, and the remaining nine students scored correctly for the two questions on properties of sides, but failed both questions on angles. This again supports the notion that angles are a more difficult concept than sides. Incorrect responses to Item 21 included the halving of ∠BAC (seven students), and the giving of a linear measurement, e.g., 10cm, for ∠EDF (six students). One student, for the question parts testing for knowledge of properties of sides, stated that she measured ED with a ruler, and found it to be 2 cm long.

S57, while unable to answer correctly the questions on angle properties, gave sufficient correct responses to achieve the criterion for Level 3, thus registering a response pattern error between the two levels. Her responses to Item 21 were (c) *30°*, and (d) *1/2 of ∠A*. All her correct Level 3 answers were for questions about either sides or general figures. Questions concerning angles were unanswered.

For the concept similarity, 43% of the students were unable to reach criteria for Level 2 or higher, a greater proportion than for other concepts. Again, students displayed a greater familiarity with properties of sides than with properties of angles. Six students (20%), in

giving a linear measurement for the angle, showed they were unaware of angular measurement.

**Level 3**

Mayberry designed thirteen question parts, with a criterion requirement of eight correct responses, to test for Level 3 skills. The question parts occur in Items 40, 41, 42 and 44, testing whether students understand conditions necessary for a definition (one part), relationships between properties (eleven parts), and class inclusion (one part). Only four students reached Mayberry's Level 3 criterion. A fifth student demonstrated understanding of relationships in his Level 4 proof for Item 58. His failure to achieve Mayberry's Level 3 criterion caused a response pattern error.

As mentioned in the previous section, Items 41 to 44 require students to work with generalised figures, no diagrams being given. Overall, although most students drew diagrams, these items were not well answered, many students failing to offer reasons. Eleven of the thirteen question parts testing for Level 3 understanding of similarity are contained in the two items, Item 41 which specifies triangles ABC and DEF, and Item 42 which refers to general figures. Of the twelve students who gave reasons for their answers to Item 41, six were considered to have given acceptable responses to at least half the question parts. For Item 42, only eight students offered any reasoning. Of these, five were considered to have given acceptable responses to three or more question parts.

S53, demonstrated ability to relate properties in his solution to Item 58 (Level 4). He did not attempt to give any reasons for Items 41 to 44, and hence, failed to reach Level 3 criterion, registering a response pattern error between Levels 3 and 4.

Overall, the results for Level 3 show that for questions about generalised figures, as in Items 41, 42 and 44, although most students drew diagrams of specific figures, more than half of the students could not answer these types of questions.

**Level 4**

A single question requiring understanding of the construction of a proof, tests for Level 4 skills, for the concept similarity. Consequently, the criterion for the level and concept is 100%. Two students were credited with Level 4 understanding.

The question designed by Mayberry to test for Level 4 understanding of the concept similarity, Item 48, has more than one proof, and, hence, can be solved without referring to

similarity. For Paper II in the preliminary study, this resulted in only one student using a proof involving similarity. No other student was able to be assessed for Level 4 understanding of similarity. Hence, as detailed in Chapter 3, a new question, Item 58, was developed. In the main study, three students attempted to construct a proof using similarity for Item 48. None completed an acceptable proof. However, it would not have been possible to evaluate the remaining twenty-seven students for their understanding at Level 4 for the concept similarity, had not the new question been developed.

Item 58 (main study)

| |
|---|
| Figures MNO and PQR are right triangles with N and Q being a right angles. MO and PR are in the ratio a:b.<br><br>What is the least additional information needed to ensure that triangles MNO and PQR are similar? |

While considered as meeting Mayberry's (1981, pp.48-49) behavioural definition for Level 4 in that each of the two students (S53 and S58) demonstrated understanding of necessary and sufficient conditions, neither of the successful students gave a generalised response for this question.

S53 *Either MN and PQ or NO and QR are also in the ratio of a:b.*

S58 *Another angle being equal, e.g.,* $\angle MON = \angle PRQ$.

Responses to this item were found to be difficult to assess. For example, a response "$\angle M = \angle P$ (corresponding angles)" can be a student's genuine recognition of the necessary and sufficient condition that a second pair of corresponding angles must be equal (Level 4), or it can be the only relationship a student can recognise in connection with the triangles (Level 3). To overcome this difficulty, the student's overall responses for the concept similarity were considered in the interpretation of the level of understanding displayed in their responses to Item 58. As a consequence, an amendment to the original question, changing from specific triangles to general triangles is offered:

Item 58 (amended)

| |
|---|
| What is the least information needed to ensure that a pair of right triangles are similar? |

This amendment does not provide prompts. An acceptable Level 4 answer would be given as a general statement; for example, "one pair of (acute) angles must be equal", or again, "two pairs of corresponding sides must be in the same ratio." The difficulty the researcher has

experienced in devising a suitable and valid item highlights the success Mayberry displays in that so many of her items are a valid test for the levels for which they are designed.

**Summary and implications**

Responses for Level 2 re-enforce the notion that understanding of properties of angles requires, in part, the acceptance of a unique measurement system. Six students (20%) showed they were unaware of the angular measurement, the degree, when they gave a size for ∠EDF in centimetres. Overall, the results show that, although all students could recognise similar figures, not many were able to demonstrate further understanding of the concept.

## Overall Results

In the previous seven sections each concept was considered separately. This section takes an integrated view by considering the results across the sections, and compares the students' performances with their mathematics background. The percentages of students achieving Mayberry's criterion for each concept and level are given in Table 4-10, and the results tabled in each section above for the highest van Hiele level achieved by the students, are summarised in Table 4-11.

Table 4-10

Students Successful at Each van Hiele Level for Each Concept

| Concept (n) | Level 1 | Level 2 | Level 3 | Level 4 |
|---|---|---|---|---|
| Square (61) | 61(100%) | 59(97%) | 5(9%) | 4(7%) |
| Right triangle (31) | 30(97%) | 23(74%) | 7(23%) | 1(3%) |
| Isosceles triangle (30) | 28(93%) | 19(63%) | 7(23%) | 1(3%) |
| Circle (31) | 31(100%) | 24(77%) | 21(68%) | 5(16%) |
| Parallel lines (30) | 30(100%) | 25(83%) | 0(0%) | 1(3%) |
| Congruency (31) | 31(100%) | 20(65%) | 3(10%) | 9(29%) |
| Similarity (30) | 30(100%) | 16(55%) | 4(13%) | 2(7%) |

Table 4-11

Highest Level Reached by Students for Each Concept

| Concept (n) | No Level | Level 1 | Level 2 | Level 3 | Level 4 |
|---|---|---|---|---|---|
| Square (61) | 0(0%) | 2(3%) | 51(84)% | 4(6.5%) | 4(6.5%) |
| Right Triangle (31) | 1(3%) | 6(20%) | 17(54%) | 6(20%) | 1(3%) |
| Isosceles Triangle (30) | 2(7%) | 8(27%) | 13(43%) | 6(20%) | 1(3%) |
| Circle (31) | 0(0%) | 4(13%) | 6(19%) | 16(52%) | 5(16%) |
| Parallel Lines (30) | 0(0%) | 5(17%) | 24(80%) | 0(0%) | 1(3%) |
| Congruency (31) | 0(0%) | 10(32%) | 11(36%) | 1(3%) | 9(29%) |
| Similarity (30) | 0(0%) | 13(43%) | 12(40%) | 3(10%) | 2(7%) |

The tables show that for the concept circle, students were more able to achieve higher level responses. The most difficult concepts were parallel lines and similarity. All students showed mastery of at least Level 1 for all concepts except for the two triangle concepts. With the exception of the circle questions, less than one third of students were able to display van Hiele Level 3 skills, i.e., two-thirds of the students could not do more than list properties associated with each concept (Level 2). Interviews supported the levels determined in the analysis of the responses to the written test. Prompting did not cause students to demonstrate a higher level of understanding than that shown in their written answers, nor was there any indication of prompting having a 'level reducing' effect as was the concern of Pegg and Davey (1989, p.19).

Mayberry used the Guttman scalogram analysis as part of her investigation into whether the van Hiele levels as tested by her questions, are hierarchical. Her results gave a coefficient of reliability of 0.95. If the translation of the Mayberry interview items to a written format has not affected the essence of the Mayberry test, then a Guttman scalogram analysis of the results should give a value for the coefficient of reproducibility similar to the value determined for the Mayberry study. Each of the sixty-one students was assessed for four van Hiele levels, across four concepts, giving a total of 976 responses. Nineteen error patterns occurred in the 976 responses, thirteen in Paper 1 and six in Paper 2.

$$\begin{aligned}\text{Coefficient of reliability} &= 1 - \text{total number of errors/total number of responses}\\ &= 1 - 19/976\\ &= 0.98\end{aligned}$$

A Guttman coefficient of reproducibility of 0.98 compares favourably with the Mayberry result of 0.95, indicating that the written test is a valid transformation of the Mayberry interview items. As recorded in the relevant sections above, the values of the Guttman coefficient of reliability for each concept, were all of an acceptable value, being 0.94 or higher (Torgenson 1967, p.323).

To make a meaningful comparison between the students' geometric backgrounds and their results, a most common level of reasoning was calculated for each student. In general, this was the van Hiele level most commonly reflected in each student's results across the four concepts. However, some students' results showed more than one common level, for example, Level 1 responses given for two concepts, and Level 2 responses for the other two concepts. For these, the most common level of understanding was taken as the higher level. In the given example, the most common level would be listed as Level 2. No student was consistently unable to achieve Level 1 criteria. The most common working levels for the students are summarised in Table 4-12.

Table 4-12

Students' Most Common van Hiele Level (n=61)

| van Hiele level | Level 1 | Level 2 | Level 3 | Level 4 |
|---|---|---|---|---|
| Number (percentage) of students | 5(8%) | 42(69%) | 6(10%) | 8(13%) |

The indication in Table 4-11 that most of the of the students were unable to demonstrate reasoning greater than van Hiele Level 2, i.e., they were unable to show the relationships between properties, is supported in the above table which shows that 77% of the students were most commonly working at Levels 1 and 2.

Fifty-five students volunteered information about their mathematical background. This information was summarised into three categories, those whose senior secondary mathematics course included a recognised geometric segment, those who completed a senior secondary mathematics course which did not contain a recognised geometry strand, and those not completing a senior secondary mathematics course. A comparison of the mathematical background for the fifty-five students with their most common van Hiele working level is shown in Table 4-13.

Table 4-13

Comparison Between Students' Geometric Backgrounds and Their Most Common van Hiele Working Level

| Geometric Background | van Hiele Level 1 | van Hiele Level 2 | van Hiele Level 3 | van Hiele Level 4 |
|---|---|---|---|---|
| Senior geometry (n = 35) | 0(0%) | 22(63%) | 5(14%) | 8(23%) |
| Senior maths but no geometry (n = 13) | 1(8%) | 12(92%) | 0(0%) | 0(0%) |
| No senior maths (n = 7) | 3(43%) | 3(43%) | 1(14%) | 0(0%) |

The findings are predictable, with the students who have completed geometry courses performing better than those with none. However, the surprise is the relatively small differences in understanding displayed among the groups. It is assumed that instruction for senior secondary courses which include a formal or recognised geometry segment is at van Hiele Level 3 or higher. Although students who have completed a senior geometry course performed better than those who have no senior geometry, it is noteworthy that twenty-two of the thirty-five students in this category were not able to demonstrate Level 3 understanding of the relationships between properties.

## Conclusion

The results tabled in this chapter show how a sample of Australian students training to be primary teachers, performed on a written test version of the Mayberry questions. They indicate also, that the Mayberry items transferred successfully from an interview situation to a written paper. This is born out in two ways; first, the levels of thinking displayed by the students in the interviews supported the levels obtained in the written test, and second, the coefficients of reproducibility calculated in a Guttman scalogram analysis compared favourably with the values calculated for the Mayberry results.

In analysing the students' responses to the test items, various patterns emerged. Many of these patterns occurred also in the Mayberry results, and were documented in her data summary (1981, p.89). The overall responses to individual questions or to clusters of questions indicated that :

1. Some students had difficulty in recognising a figure in a non-standard position.
2. Properties involving sides of figures were more readily identified than properties involving angles.

3. Difficulties arose for some students in establishing the context for questions designed to test Levels 1 and 2.
4. Many students did not have an understanding of class inclusion, for example, they did not recognise a square as a rectangle.
5. Students had no understanding of the meaning and/or significance of necessary and sufficient conditions.
6. Some students, working at Level 2, resorted to measuring sides or allocating values to angles when presented with a Level 3 problem.
7. Students who were unable to work with a generalised figure, usually drew a geometric figure.
8. The highest level reached by more than half of the students who had completed a senior secondary geometry course was Level 2.

The next chapter continues the analysis of the students' performances on the test. In particular, the focus is on a quantitative approach where the students' responses are subjected to an analysis using the QUEST program.