CHAPTER 3: METHODOLOGY

3.1 Research Design

This study adopted a quasi-experimental design using an intact class of Form One students as the subjects of the study. The defining characteristic of a quasi-experimental design is that it employs groups that are already formed so that individuals are not randomly assigned to treatment conditions (McGuigan, 1990). McGuigan asserts that quasi-experimental designs are less rigourous than true experimental designs, which yield results with the highest probability. However, in order to improve educational practices, "we should accumulate as much knowledge of as high degree of probability as we can. For such a purpose we *need* quasi-experimental designs" (p. 273).

More specifically, an untreated comparison-group design with pretest-posttest was adopted for this study. Such a design first established the 'equality' of the treatment and the comparison groups through the administration of a pretest. The pretest also established a baseline for the treatment group prior to the intervention with the treatment. A posttest then measured the efficacy of the treatment. The posttest was also administered to the comparison group for comparison purposes (McGuigan, 1990; Norusis, 1991).

In the present study, experimenter-designed pre- and posttests were used. The treatment condition consisted of instruction in drawing inferences of four different types. The instruction provided was grounded in a direct instruction paradigm.

3.2 Subjects

An intact Form One class of forty-four students in Anderson Secondary School, Ipoh was used for this study. The selection of an intact group as the subjects of this study was subject to Hatch and Lazaraton's (1991) observation that the "majority of classroom research involves the use of classes where students have already been assigned on the basis of some principle" (p. 85). In Anderson Secondary School, first form students are streamed on the basis of their overall performance in three subjects (Bahasa Malaysia, English, and Mathematics) in the Primary School Assessment Examination at the end of Year Six of the primary school programme.

A pool of three classes of similar ability was first identified based on the ratings of the English Language teachers and the Subject Head for English Language, and on the performance of the classes on the monthly and first semester exams. From this pool, one class was randomly selected to be the treatment group for this study.

The treatment class, which had been randomly selected from the pool of three classes, had forty-four students. However, six of the students in the treatment class missed one or more of the measures that were administered. Consequently, these six students were discarded from the analysis of the findings of this study. Thus, the treatment class comprised thirty-eight subjects.

Anderson Secondary School is an all boys' school and the treatment group comprised all three major racial groups - Malays, Chinese and Indians. The ethnic background of the treatment is shown in Table 3.1 below.

Table 3.1 Ethnic background of the treatment group

Ethnicity	Number	Percentage
Malay	26	68.4
Chinese	6	15.8
Indian	6	15.8
Total	38	100

An English Language Test devised for the Profiles of the Development of Primary School Children (Kaur, 1996) was then administered to the treatment group in order to determine the English Language ability of the students and to categorise the subjects in the treatment group according to their English Language proficiency. This test will be referred to as the Assessment Test in this study. The Assessment Test comprised three components: vocabulary, reading and grammar.

On the basis of their performance on the Assessment Test, those subjects in the Treatment group who scored below the 33rd percentile were designated as Low English Language Proficiency students while those who scored above the 67th percentile were designated as High English Language Proficency students. The designation of the subjects in this manner was confirmed against the ratings of the English teacher teaching the class.

Table 3.2 below shows the mean scores of the three English Language ability groups in the treatment group that were identified on the basis of their performance on the Assessment Test. As can be seen in Table 3.2 below, the

mean scores of the three groups differed quite substantially, with the high ability group having a mean of 69.17, whereas the median and low ability groups had mean scores of 47.07 and 29.5 respectively.

Table 3.2 Mean scores of the three English Language ability groups in the treatment group on the Assessment Test

Ability Group	Number	%age	Mean	s.d.
High Proficiency	12	31.6	69.17	11.25
Median Proficiency	14	36.8	47.07	6.59
Low Proficiency	12	31.6	29.50	6.49
Total	38	100		

In order to determine if a significant difference existed between these three groups in their English Language proficiency, as measured by the Assessment Test, a Oneway between-groups ANOVA was carried out (Table 3.3 below). The independent variable in this analysis was the three-level grouping of the subjects, and the dependent variable was their performance on the Assessment Test.

Table 3.3 Oneway ANOVA for the High, Median and Low Proficiency subjects from the treatment group on the Assessment Test

Source of variance	*df*	Sum of squares	Mean squares	*F* Ratio	*F* Prob
Between groups	2	9485.9048	4742.9524	68.5512	.0000
Within groups	35	2421.5952	69.1884		
Total	37	11907.5000			

Scheffe Procedure

Ranges for the .050 level

3.62 3.62

The value actually compared with Mean (J) - Mean (I) is

5.8817 * Range * Sqrt(1/N(I) + 1/N(J)

Mean	Group	Low	Median	High
29.50	Low			
47.07	Median	*		
69.17	High	*	*	

* Denotes pairs of groups significantly different at the .050 level

The Oneway ANOVA yielded $F = 68.55$ (2, 35), $p < .05$, which indicated that there was a difference between the means of the three ability groups. In order to determine which of the three groups differed from one another, the Scheffe Test was carried out. Among the various multiple comparison tests available, the Scheffe procedure is considered to be most conservative since it requires larger differences between means for significance than most other methods (Norusis, 1991). The Scheffe procedure indicated that the High, Median and Low English Language Proficiency groups

in the treatment group differed significantly from each other in terms of their English Language Proficiency as measured on the Assessment Test.

It should be noted that this designation of the subjects of the study into high, median, and low English Language proficiency was carried out for data analysis purposes only. All the students continued to remain a part of the class and underwent the same instruction together as an intact group.

A comparison class of forty-two students was also identified. Like the treatment group, the comparison group comprised an intact class, and was randomly selected from the pool of three classes that had been identified earlier.

Five of the students in the comparison group missed one or more of the measures that were administered in the course of the study and they were thus excluded from the data analysis in this study. The comparison group thus comprised thirty-seven students. The comparison group, like the treatment group was also divided into three groups based on their English Language ability as measured on the Assessment Test (Table 3.4).

Table 3.4 Mean scores of the three English Language ability groups in the comparison group on the Assessment Test

Ability Group	Number	%age	Mean	s.d.
High Proficiency	12	32.4	68.13	11.92
Median Proficiency	13	35.2	45.19	5.72
Low Proficiency	12	32.4	28.96	4.82
Total	37	100		

A Oneway between-groups ANOVA was carried out (Table 3.5 below) in order to determine if a significant difference existed between these three groups in their English Language proficiency, as measured by the Assessment Test.

Table 3.5 Oneway ANOVA for the High, Median and Low Proficiency subjects from the comparison group on the Assessment Test

Source of variance	*df*	Sum of squares	Mean squares	*F* Ratio	*F* Prob
Between groups	2	9298.7634	4649.3817	71.4300	.0000
Within groups	35	2213.0609	69.0900		
Total	37	11511.8243			

Scheffe Procedure

Ranges for the .050 level

3.62 3.62

The value actually compared with Mean (J) - Mean (I) is

5.7048 * Range * Sqrt(1/N(I) + 1/N(J)

Mean	Group	High	Median	Low
68.13	High			
45.19	Median	*		
28.96	Low	*	*	

* Denotes pairs of groups significantly different at the .050 level

The Oneway ANOVA yielded $F = 71.43$ (2, 35), $p < .05$, which indicated that there was a difference between the means of the three ability groups. Additionally, the Scheffe procedure indicated that the High, Median

and Low English Language Proficiency groups in the comparison group differed significantly from each other in terms of their English Language Proficiency as measured on the Assessment Test.

Finally, a two sample *t*-test was carried out to determine if the comparison and treatment groups differed in terms of their English Language ability as measured on the Assessment Test. The findings are displayed in Table 3.6 below.

Table 3.6 Performance of the comparison and treatment groups on the Assessment Test

Group	*n*	Mean	*s*	*t* value	*df*	*p*
Comparison	37	47.36	17.88	-.27	73	.785
Treatment	38	48.50	17.931			

The comparison group had a mean of 47.36 with a standard deviation of 17.88 whereas the treatment group had a slightly higher mean of 48.50 and a standard devistion of 17.931. However, the *t*-test yielded t = -.27, df = 73, p = n.s., thus indicating that the comparison and the treatment groups were not significantly different in terms of their English Language ability as measured on the Assessment Test.

3.3 The Pre- and Posttests

In addition to the Assessment Test, experimenter-designed pre- and posttests were also used in this study.

The pretest used in the study was administered immediately prior to the provision of instruction in inference making. The pretest served two purposes. First, it established the 'equality' of the treatment and the comparison groups. Second, the pretest also established a baseline for comparison of the subjects' performance on the posttest.

The pre- and posttests tested the inference making ability of the subjects of the study. They comprised items which tested inference types based on Chikalanga's (1993) categories of inferencing: i) logical informational; ii) logical explanatory; iii) elaborative informational; and, iv) elaborative explanatory inferences.

The pre- and posttests each comprised four narrative passages. Five multiple choice (MCQs) items were formulated for each passage, thus resulting in a total of twenty MCQ items for each test. The five questions for each passage assessed at least one of the four categories of inferences stated above. For both the pretest and posttest, as a whole, therefore, there were five questions testing each one of the four categories of inferences. Following the recommendation made by Sundbye (1987) and Wilson (1979), the sequence of the multiple-choice items for each passage followed the sequence of key events in the respective narrative texts.

Multiple-choice items were preferred in the tests to open-ended questions because of the subjects' limited exposure to English. It was anticipated that the subjects might have difficulty expressing themselves and

would, therefore, be unable to give an account of their true inference making ability if the open-ended format had been used. When answering MCQs, the subjects needed only to recognize the key to each item although the cognitive demand was the same. That is, they still had to fall back on their prior knowledge, make use of contextual clues, and integrate information across sentences or paragraphs in order to identify the correct answer.

The posttest was administered at the conclusion of the training phase of this study. For the pre- and posttest, no time limit was set for the subjects to complete the questions.

3.4 The Trialling of the Tests

The pre- and posttests were tried out in a different school from that used in the study. The tests were administered to a Form One class of 36 students. The tests were administered on consecutive days.

The analysis of the data derived from the trialling of the tests included (a) establishing the internal consistency of each of the tests by computing the Cronbach *alpha*; and (b) performing an item analysis of all the items comprising each of the tests. The QUEST Interactive Test Analysis System software programme (Adams and Khoo, 1993) was used to complete these two procedures.

The mean scores and standard deviations obtained in the pilot testing as well as the Cronbach *alpha* obtained for each of the tests is shown in Table 3.7 below.

Table 3.7 Means, standard deviations and Cronbach *alpha* for the pre- and posttests

	Mean	s.d.	*alpha*
Pretest	11.45	2.95	.77
Posttest	10.81	2.81	.80

Of the two tests, the pretest had the lower internal consistency of .77. The item analysis further indicated that the largest number of problem items was in the pretest. The posttest had an internal consistency of .80. The performance of the pilot test subjects on the two tests was then correlated and is shown in the correlation matrix in Table 3.8 below.

Table 3.8 Correlation matrix of the performance of the pilot test subjects on the pre- and posttests

	Pretest	Posttest	All
Pretest	-	.76	.85
Posttest	-	-	.88

The correlation matrix above suggests that there was a fairly high correlation between the performance of the subjects on the two tests. A correlation coefficent of .76 for the pre- and posttests was obtained. Both tests correlated very well with the subjects' performance on the combined score on the two tests, with the pretest having a correlation coefficient of .85, and the posttest having a correlation coefficient of .88.

3.5 Instructional Design

This study adopted a direct instruction paradigm to teach the subjects how to draw inferences from textually implicit and scriptally implicit information. Many researchers have made claims on behalf of the effectiveness and usefulness of a direct instructional paradigm in teaching reading comprehension strategies. Baumann (1984b), for example, made it clear that he had adopted a direct instructional paradigm to teach his subjects main idea comprehension because "of the success researchers have had in training students in comprehension skills when instruction has been systematic and direct" (p. 95).

The present study deployed the direct instruction paradigm for the provision of instruction in inference making. Four types of inferences, from Chikalanga's (1993) taxonomy of inferences, were the focus of the instruction provided in this study. The four inference types were: a) logical informational inferences (that is, inferences that require students to determine the people, things, time, place and general context of given events; b) logical explanatory inferences (that is, inferences that determine motivations of characters, causes and consequences of events and actions stated in the text and the conditions which enable events and actions to occur); c) elaborative informational inferences; and, d) elaborative explantory inferences. The last two categories of inferences are similar to logical informational and logical explanatory inferences except that the elaborative inferences are based on information "outside the text". That is, the reader has to rely on her schema in order to draw appropriate inferences (Chikalanga, 1993).

3.5.1 Theoretical Basis of the Training Procedure

The training was carried out by the experimenter in an intact First Form class. Instruction was provided during normal class hours alloted to the teaching of English. Instruction was provided for 2½ hours per week for a duration of five weeks.

The training was grounded in a direct instruction paradigm. Baumann's (1988) and Duffy & Roehler's (1982) definition of direct instruction mandated the following key instructional procedures: a) deciding on the goal of each lesson (academic focus); b) sequencing the content of each lesson; c) modeling, demonstrating, and providing explanations related to the strategy; d) providing guided practice and giving feedback to students' responses; and, e) ensuring that there is a gradual shift of responsibility in the use of the new strategy from the teacher to the students.

Specifically, in adopting a direct instructional paradigm, Baumann & Schmitt's "What, Why, How, and When" instructional model (in Baumann,1991) formed the basis of instruction. This model was preferred because it encompasses all the important characteristics of direct instruction.

First, the model recognizes the importance of not only the "what" (declarative knowledge), but also the "how" (procedural knowledge), "why", and "when" (conditional knowledge) of the strategy in which instruction is being provided. Simpson (1985) asserts that students should not merely be taught the steps involved in the use of the new strategy (the "what and how"). Equally important is the metacognitive information related to the strategy (the "why and when").

Simpson (1985) uses the term "blind training" to refer to instruction where teachers do not inform their students of these aspects of strategy use. In contrast to "blind training" is instruction which is termed as "informed training" where students are told about the "what, how, why and when" of the strategy (Oxford, 1989). Knowledge of the "why and when" aspects is important because it enables students to use the strategy flexibly and appropriately.

Second, the model emphasizes direct and explicit teacher explanations, modeling and demonstration of the strategy. To do this, the teacher has to assume an overt instructional role as a means of achieving process and content goals. Through "verbal mediation", the teacher directly and explicitly explains to the students how to use the strategy being taught, saying aloud the mental steps she goes through when using the strategy.

An example of this would be: "When I use this strategy, I ought to read the first two or three sentences. Then I need to decide what information the author has left unstated. Then I..." (adapted from Duffy & Roehler, 1989, p. 231). Duffy and Roehler assert that "making visible the thinking that leads to the answer" (p. 231) is helpful to the students and that this "mental modeling" provides the students with a model to follow when they try the new strategy. Graves, Watts, & Graves (1994) assert that "modeling is a window on the mind and certainly one of the most poweful tools for showing children how to think" (p.43).

Third, this model emphasizes the need for guided practice. In this "interactive phase" of the lesson, the teacher mediates the students' acquisition of the strategy. In the initial stage of guided practice, the teacher closely monitors and guides the students as they practise using the strategy. In

so doing, the teacher gradually moves the students to the point where they can use the strategy independently. This stage of the lesson thus sees the teacher gradually withdrawing assistance as the students exhibit success in using the strategy.

The thrust of the lesson at this point is, therefore, to produce a gradual transition from teacher modeling to student control of the use of the strategy. In order to effect this transfer, Duffy and Roehler (1989) caution that the lesson must be carried out at a pace appropriate to the students.

At this point, too, the successful monitoring of the students' progress in the use of the strategy depends, to a large extent, on the activities assigned to the students. Duffy & Roehler (1989) assert that since the focus of the instruction is the teaching of a reading strategy, the teacher should be more interested in the thinking processes of the students in arriving at the answer rather than in the accuracy of the answer that they give. The kinds of questions posed to the students should, therefore, be those that monitor the students' thinking processes. For example, questions like "How do you arrive at this [inference]?" should be asked rather than "What do you [infer]?" (p.234). The second question is more an assessment of how much the students know as well as how accurate their answer is. On the other hand, the first question is related to the processes by which the answer is derived.

The key points of Baumann & Schmitt's "What, Why, How, and When" instructional model (in Baumann, 1991) which have been discussed above formed the basis of the training procedure, which is detailed in the next section.

3.5.2 The Training Procedure

The key steps in the training procedure which was used to provide instruction in inference making are detailed below. The training procedure was based on the Baumann and Schmitt "What, Why, How and When" instructional model (in Baumann, 1991) discussed above.

Step 1: **What** Introduce the strategy to the students. Explain that writers often leave out many facts and ideas when they write. Tell the students that it is their job to figure out those unstated ideas.

Step 2: **Why** Explain why the strategy is important.

Step 3: **How** First, based on Chikalanga's (1993) logical informational category of inferences (see Figure 3), discuss with the students the different types of unstated information in texts.

Second, explain and model the strategy used when inferring unstated information by using a "How to figure out unstated ideas" chart, (see Figure 4, from Baumann, 1991, p.68). For this purpose, a short familiar story `The Pencil' will be used.

Third, get the students to work in pairs / groups and try out the strategy with another story.

Step 4: **When** Remind the students that when they read, they need to understand what the writer explicitly states in the text. Then they need to figure out what ideas / facts the writer has left out. They need to draw inferences when reading all kinds of reading materials- stories, school-related texts and newspapers because this will help them to understand what they read better.

Figure 3. Instructional Chart Listing Six Inference Categories

UNSTATED IDEAS

1. **PEOPLE OR ANIMALS**
Polly was up to bat. Her team was down by two runs. She swung at the pitch and hit a homer.

Who is Polly?

The farmer fed table scraps to the squealing animals. They rooted in the feeding trough with their noses and made snorting, oinking sounds as they ate.

What are the animals?

2. **PLACES**
Max took out his pencil and math paper when the teacher told the class to get to work.

Where is Max?

3. **THINGS**
Kristi pushed the roaring machine all over the lawn. When she was finished, the grass looked beautiful.

What is the roaring machine?

4. **TIME**

The bell rang, and Ms. Olson said, "All right children. Go outside and play hard for 20 minutes, but be ready to take your spelling test when you come back inside."

What time was it?

5. **ACTIONS**

Jim took an end and let Nancy jump for a while. We all sang as she jumped. Nancy finally missed, and then she took an end.

What were Nancy and Jim doing?

6. **FEELINGS**

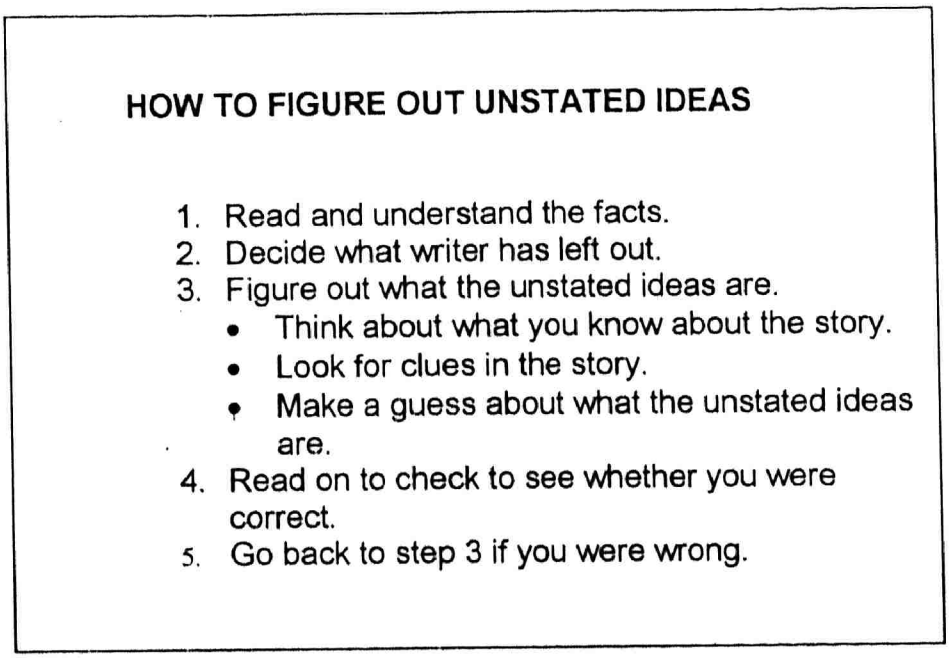
After his team lost the kickball game in gym class, Scott walked over to the bleachers, kicked them with his foot, and yelled, "darn".

How was Scott feeling?

(from: Baumann, 1991; p.67)

A complete script for a sample lesson based on this four-step training procedure is appended (Appendix 1).

Figure 4. A strategy for inferring unstated information



(Adapted from *Making Inferences* by Dale D. Johnson and Bonnie von Holf Johnson, cited in Baumann, 1991)

Finally, it must be emphasized that although the lesson steps for each lesson will be pre-planned, there will still be room for some degree of flexibility in the actual delivery of the lesson in the classroom. In this regard Duffy & Roehler's (1989) caution about prepared lesson formats is worth noting. They assert that though a teacher must be equipped with lesson formats, she must, however, be ready at all times to make "spontanteous subtle distinctions and adjustments as she teaches" (p. 235).

3.5.3 Instructional Materials

The instructional materials used in this study were based on narrative texts. A narrative text is defined as a stuctural form of writing that relates a story in either prose or verse (Harris & Hodges in Cooter & Reutzel, 1994)

It was decided that narratives would be used in this study because most children have either heard / read or have had stories read / told to them from a very young age. They are, therefore, more familiar with narratives and their organizational structures than they are with the organizational structures of expository texts (Cooter & Reutzel, 1994; Faman, Flood, & Lapp, 1994). Thus, by using narrative texts, it was anticipated that the confounding effects of text structure on the findings of this study would be eliminated.

Additionally, Richek, List, & Lerner (1989), in their comparison of expository and narrative texts, highlight the fact that the former tend to be more difficult than narrative texts because they tend a) to consist of more difficult words; b) are usually much longer; and, c) they tend to be more "inconsiderate" (that is, they are less systematic in terms of organization) thus making them more difficult to read and understand than are narrative texts. This was an important consideration given the subjects' limited exposure to English.

The narrative texts used were also culturally familiar to the subjects. This decision was influenced by Chikalanga's (1993) finding that students are better able to draw inferences based on texts which are culturally familiar. Chikalanga suggests that ESL teachers should use reading materials that are appropriate to the "cultural" or "domain-related" schemata of their students. The use of culturally familiar texts helped to ensure that content cultural schema did not confound the findings.

3.6 Data Analysis

The primary data for the study was obtained from the high and low proficiency subjects' performance on the Pretest, and the Posttest, after the provision of direct instruction in inference making.

Descriptive statistics (means, percentages, and standard deviations) were used to provide an overview of the Treatment and Comparison groups' performance on the Pretest and the Posttest. Between-groups *t*-tests were computed to determine whether the Treatment and Comparison groups' performance on the Pretest and Posttest were significantly different.

A between-groups ANOVA for a 2 x 2 design was computed to determine the effect of the treatment condition and the language proficiency of the subjects on their performance in inference making on the Pretest and Posttest.

Finally, means and standard deviations were used to provide an overview of the High and Low Proficiency subjects on the four inference types. Between-groups *t*-tests were computed to determine whether the performance of the High and Low Proficiency subjects was significantly different on the Pretest and the Posttest.

The data was analyzed with the Statistical Package for the Social Sciences (SPSS) software.