CHAPTER 4

RESEARCH METHODOLOGY

4.1 Research problem

4.1.1 Aim of the research

The aim of this research is to design and develop a computer-assisted tutorial "Random Variables" and then to investigate and evaluate the impact of this program on a sample of students from the target population.

4.1.2 Objectives of the research

In order to achieve the above aim, the researcher's objectives were to:

- analyse the target population, the subject matter and appropriate hardware and software;
- design the tutorial "Random Variables" according to an instructional strategy incorporating cognitive learning theories;
- develop and test the tutorial "Random Variables" using the TenCore authoring package;
- modify the prototype after peer review by subject matter and computer-assisted instruction experts;

- implement the six tutorials, which have been developed to date, for students to work through on their own computers or in the Unisa regional micro-computer laboratories;

- evaluate the tutorial “Random Variables” by obtaining peer reviews and student feedback about various aspects of the program;

- assess what changes and modifications are required to improve the tutorial “Random Variables”.

4.2 Research questions

In order to evaluate the impact of the tutorial “Random Variables”, the following main questions arise:

1. What corrections and modifications to the program are required?
2. What are the cultural and language implications for the heterogeneous student population?
3. Does the teaching approach embodied in the program contribute to perceived learning gains?
4. Is it clear how to use the function keys and icons to navigate through the program?
5. What are the opinions, feelings and emotions of the learners on completion of the program?

Each of the main questions was investigated by considering the following sub questions:

	QUESTIONS AND SUB-QUESTIONS
1.	What corrections and modifications to the program are required? 1.1 Changes to the screen displays? 1.2 Changes to text layout? 1.3 Changes to colour usage? 1.4 Typographical errors? 1.5 Content errors? 1.6 Technical hitches?
2.	What are the cultural and language implications for the heterogeneous student population? 2.1 Is anything in the programs offensive to certain cultures? 2.2 Is the level of English usage acceptable? 2.3 Is the use of vocabulary acceptable? 2.4 Is there any language which is vague or ambiguous?
3.	Does the teaching approach embodied in the program contribute to perceived learning gains? 3.1 Is intuitive understanding of basic concepts facilitated? 3.2 Is the assumed prior knowledge acceptable? 3.3 Is the program suitable for adult learners? 3.4 Are terms and concepts clearly defined? 3.5 Does the student feel that learning has taken place? 3.6 Are the questions of an acceptable difficulty level? 3.7 Is it clear how answers should be entered? 3.8 Do the question episodes and feedback reinforce learning?
4.	Is it clear how to use the function keys and icons to navigate through the program? 4.1 Is it clear how to move forwards and backwards through the program? 4.2 Is it clear how to make choices from the menus? 4.3 Is it clear how to use the function keys for special actions?
5.	What are the opinions, feelings and emotions of the learners on completion of the program? 5.1 What are the opinions, feelings, emotions of the learners? 5.2 Are the graphics helpful and the comic characters acceptable? 5.3 How do the students rate the program?

Table 4.1 Research questions and sub questions

4.3 Data collection

4.3.1 Description of methods and instruments

The instructional design model (Figure 4.1) on which this study is based, ensures that formative evaluation (and hence data collection) takes place at each of the stages of design, development and implementation.

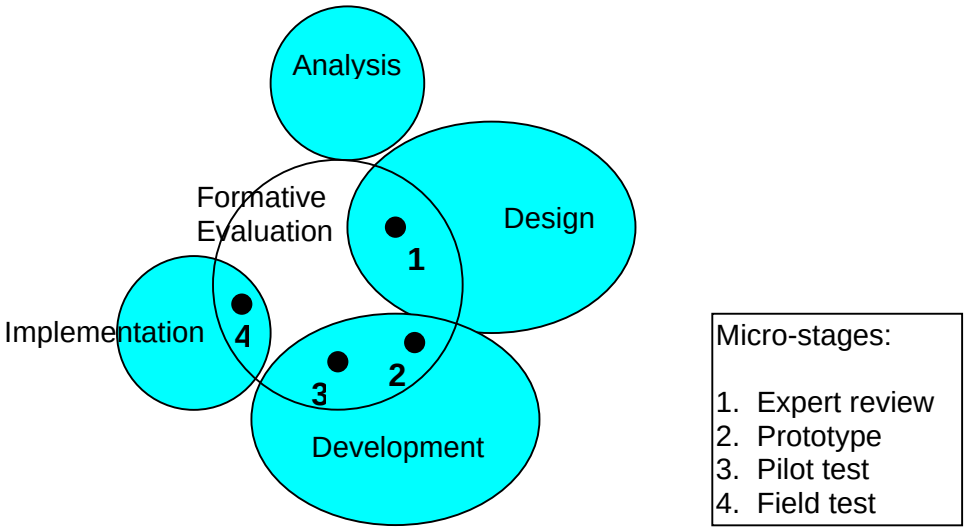

Figure 4.1 Instructional design model for this study (Adapted from Hodgkinson (1996) and Beyer (1995))

This chapter discusses the formative evaluation that took place at each micro-stage in the instructional design process.

Denzin (cited in Mouton & Marais, 1993) coined the term 'triangulation' to refer to the use of multiple data sources, multiple methods and multiple perspectives. According to Mouton and Marais (1993), triangulation is likely to increase reliability, since the complementary nature of multiple methods can counteract their respective shortcomings.

In order to increase the reliability of this study through triangulation, several different data collection methods and instruments were used. These are summarised according to the instructional design model, in Table 4.2.

STAGE	MICRO-STAGE	METHOD	INSTRUMENT(S)
Design	1. Expert review	Peer review	Evaluation notes
Development	2. Prototype	Peer review Focus groups	Evaluation notes
Development	3. Pilot test	Survey Discussion	Questionnaire Error log Anecdotal records
Implementation	4. Field test	Survey Observation Telephone interviews	Questionnaire Error log Field notes Interview schedule

Table 4.2 Methods and instruments

4.3.2 Data analysis methods

It must be remembered that this study is intended to evaluate the product and not to measure the amount of learning that took place. As such, it is largely qualitative and descriptive in nature, with little statistical analysis, beyond frequency counts and the graphical representation of data.

The binomial test was used as a guideline for deciding when the frequency of a particular response to the questionnaire was statistically significant. According to binomial tables for a sample size of n=25, a frequency of 18 (i.e. 72%) is statistically significant (Hollander & Wolfe, 1973). For n=25, this statistic is the same for a two-sided and a one-sided test.

4.3.3 Data collection matrix

The data collection matrix, which shows which methods were used to collect data with respect to each main research question, as well as the team members involved at each stage, is given in Table 4.3.

Method / Question	Peer review Focus groups	Survey	Observation	Telephone Interviews
Corrections/ modifications?	SME, CAIE, R	L SME	R	
Cultural / language implications?	SME, CAIE, R	L		
Teaching approach?	SME, CAIE, R	L SME		L
Clear navigation?	SME, CAIE, R	L SME	R	
Opinions, feelings and emotions?		L SME	R	L
Data analysis method	edits to storyboards and prototype	descriptive statistics; binomial test	narrative	discussion

L = Learners
R = Researcher
SME = Subject Matter Experts
CAIE = Computer-Assisted Instruction Experts

Table 4.3 Data collection matrix

4.4 Design stage: Expert review of the design

The CATs were designed in detail on paper (storyboarded). It was necessary to produce a complete set of detailed storyboards, so that the design could be evaluated at different times by various team members.

The paper-based design was evaluated by two subject advisors in the Department of Statistics and by three CAI advisors in CENSE. Comments and suggestions were noted (evaluation notes) and then the storyboards were returned to the author for modification before authoring began.

4.5 Development stage: Evaluation of prototype

After each tutorial had been programmed and sample diskettes had been produced, the evaluation cycle continued, with the same group of reviewers thoroughly evaluating the computer-based product. In so doing, they paid particular attention to technical details such as navigation, icons etc. and to didactic details, such as when a replay option or modifications to presentation, questions and feedback were required.

Their comments and suggestions were recorded on paper and discussed in focus groups, and then agreed changes and edits were implemented by the author.

Version 2 of the prototype was then produced.

4.6 Development stage: Pilot test with staff members

A pilot test was conducted with five staff members during January 1996. The staff members worked through the tutorial “Random Variables”, as if they were students,

and then completed the draft questionnaire and the error log. The error log was two separate sheets of paper given to each participant on which they were asked to record details of any typographical errors, content errors, technical hitches or ambiguous language.

The staff members were encouraged to make suggestions as to how the questionnaire could be improved to eliminate ambiguities. These suggestions are discussed in Chapter 5, section 5.3.

The final questionnaire appears in Appendix F and the error log in Appendix G.

4.7 Implementation stage: Field test with students

4.7.1 Preparation of the questionnaire

The design and compilation of a questionnaire is not a trivial task. Special care must be taken to ensure that the questions are concise and unambiguous and that they will test what the researcher wishes to test.

Duverger (1964, p.144) states: *"The preparation of questionnaires is a complex and delicate operation. The nature, form and order of the questions is of great importance to the results of the inquiry."* He goes on to add:

> *"A questionnaire is not a series of questions in any order. It is, on the contrary, a group of questions the order of which is very carefully studied. The order, the number of questions about the same subject, their grouping in 'batteries' present difficult and important problems"* (Duverger, 1964, p.148).

In developing the questionnaire for this study, other questionnaires for the evaluation of computer-assisted lessons were consulted (Möller, 1993; de Villiers, 1993; Cronjé, 1993). Customised questions for this project were drafted and reviewed with experts in the field of questionnaire design, both at Unisa and at the University of Pretoria.

After five revisions, the questionnaire was translated into Afrikaans and both the English and Afrikaans versions were typed. Even then, it was discovered later, that in translation, the emphasis of one particular question had been unintentionally switched from positive to negative.

Reliability

Mouton and Marais (1993) describe various orientations of respondents, which may affect the reliability of a test. For example, the well-known Hawthorne effect is an example of a social desirability tendency, with the respondent attempting to give those responses which are considered desirable and admirable.

Another participant effect described by Mouton and Marais (1993) is what they call the *"acquiescent response set"* (p.889), which describes the tendency to answer "yes" or "no" to virtually all the items in a questionnaire. They refer to Sletto (cited in Mouton & Marais, 1993), who, as early as 1937, found that respondents were more likely to agree with a statement than to disagree with the inverse of that statement.

In order to counteract this tendency, several statements which anticipated disagreement were included in the questionnaire.

The table on the following page shows how each research question is addressed by one or more particular questions on the questionnaire and/or the error log.

	QUESTIONS and SUB-QUESTIONS	QUESTIONNAIRE numbers	ERROR LOG
1.	What corrections and modifications to the program are required?		
	1.1 Changes to the screen displays?	17	
	1.2 Changes to text layout?	18	
	1.3 Changes to colour usage?	19	
	1.4 Typographical errors?		yes
	1.5 Content errors?		yes
	1.6 Technical hitches?		yes
2.	What are the cultural and language implications for the heterogeneous student population?		
	2.1 Is anything in the programs offensive to certain cultures?	12, 13, 19	
	2.2 Is the level of English usage acceptable?	32	
	2.3 Is the use of vocabulary acceptable?	33	
	2.4 Is there any language which is vague or ambiguous?	34	yes
3.	Teaching approach?		
	3.1 Is intuitive understanding of basic concepts facilitated?	15, 16	
		11	yes
	3.2 Is the assumed prior knowledge acceptable?	14	
	3.3 Is the program suitable for adult learners?	35	
	3.4 Are terms and concepts clearly defined?	41,42	
	3.5 Does the student feel that learning has taken place?	38	
	3.6 Are the questions of an acceptable difficulty level?	37	
	3.7 Is it clear how answers should be entered?	36, 39, 40	
	3.8 Do the question episodes and feedback reinforce learning?		

4.	Clear how to navigation through the program? 4.1 Is it clear how to move forwards and backwards through the lesson? 4.2 Is it clear how to make choices from the menus? 4.3 Is it clear how to use the function keys for special actions?	 23, 24 25, 26, 27 28	
5.	Opinions, feelings and emotions of the learners? 5.1 Are the graphics helpful and the comic characters acceptable? 5.2 What are the opinions, feelings, emotions of the learners? 5.3 How do the students rate the program?	 20, 21, 22 43 - 49, 51 and 54 - 56 50, 52, 53	

Table 4.4 Matrix matching research questions with questionnaire and error log

4.7.2 Sampling procedure

In preparing for a field test with students, it was necessary to consider what sampling procedure to use. The difficulty with a distance education institution is that travel restrictions make it almost impossible to select a random sample of students, since students registered for the course are geographically dispersed.

It was decided to take a convenience sample from students living in Gauteng, who would be prepared to travel to the micro-computer laboratory on the Unisa campus in Pretoria. According to this method of sampling, the nearest or most convenient population elements are drawn into the sample (Cohen & Manion, 1994). As an incentive to attend, students were offered a cash contribution towards their travelling expenses. Fruit juice was provided but students were requested to supply their own meals for the day.

The students who took part in the study were volunteers and therefore did not necessarily constitute a representative sample of the student population registered for STA101-H. Cohen and Manion (1994) discuss the merits of a convenience sample, notwithstanding the fact that one cannot generalise from the findings:

> *"Small-scale surveys often resort to the use of non-probability samples because, despite the disadvantages that arise form their non-representativeness, they are far less complicated to set up, are considerably less expensive, and can prove perfectly adequate where researchers do not intend to generalize their findings beyond the sample in question"* (Cohen & Manion, 1994, p. 88).

Gonick & Smith (1993, p.97)

In order to broaden the sample, the CATs were installed in the regional offices in Cape Town and Durban and students who attended the discussion classes there during May 1996, were encouraged to work through the CATs and to complete the questionnaire.

No funding was available for the researcher to travel to Durban and Cape Town to run field tests there. There was also the logistical difficulty of asking students in those regions to attend the full day discussion class as well as to make further travel

and/or leave arrangements to attend a field test of the CATs, which would take another full day.

Unfortunately the attendance at the regional discussion classes was minimal and no completed questionnaires were received from Durban or Cape Town.

A possibility for further research would be to investigate how to obtain a representative random sample of respondents and to conduct a field test on a larger scale, perhaps offering some type of course credit to participants as an incentive. Without some sort of incentive, it is doubtful whether many students would complete and return voluntary questionnaires. Alexander et al. (1992) found that only three out of 28 students registered for a module in computer science returned voluntary questionnaires.

4.7.3 Conducting the field test

Field testing is part of formative evaluation, in that the main purpose of the field test is to determine if changes need to be made, based on use of the product by students from the target population (Callison & Haycock, 1988). These authors describe field testing as follows:

> *"Field testing involves the actual use of the materials with students and teachers interacting within the normal educational setting for which the materials have been designed for future use"* (Callison & Haycock, 1988, p.26).

The tutorial letter requesting participants for the field test was sent to all 339 students registered for STA101-H during January 1996. The field test was scheduled for March 1996. A copy of the tutorial letter is included in Appendix E.

It was decided that the students would be asked to work through two of the CATs, namely "Probability I" and "Random Variables" and that 10 hours should be allowed

for the exercise, including breaks. Three alternative arrangements were offered: a full day on a Friday, or a Friday afternoon and Saturday morning, or two consecutive Saturday mornings. Since most of the students are employed full time and study part time, it was necessary to consider constraints imposed on them by their working conditions.

Thirty five students responded to the invitation to participate in the field test. Ultimately only 25 attended and completed the questionnaire.

The field test was conducted as follows:

1. The students were welcomed and thanked for attending; they were told about the project and the availability of the CATs for purchase and home use.
2. The procedure for the day was explained, namely that they would work through two of the CATs and fill in the questionnaire on completion.
3. The questionnaire was not handed out at the start of the day, because it was necessary for the students to concentrate on the CATs and to form their opinions without being influenced by what appeared on the questionnaire.
4. The anonymity of the author as the designer and developer of the CATs was maintained, in order to avoid the Hawthorne effect (Mouton & Marais, 1993) of respondents providing replies that they might think are desirable.
5. The students were asked to look out for features in the CATs, such as screen layout, use of language, teaching approach, use of graphics, questions-answers-feedback etc.
6. Terms that might have been unfamiliar were explained, such as **graphics**, **feedback**, **icons,** etc.

One of the objectives of the study was to investigate what changes and modifications might be necessary in order to improve the tutorials. The students needed to be able to record more than what was asked for on the questionnaire. They were therefore supplied with an error log on which they were asked to record details of any

typographical errors, content errors, technical hitches or ambiguous language. The error log is reproduced in Appendix G.

The students worked through the two CATs at their own pace, each at their own computer. Supervisors were on hand to answer questions, both about the use of the computer and about the subject matter, when necessary. The researcher observed the students and made field notes about their progress, time taken, body language etc. Most students were able to complete the two CATs in a half-day session, although some did return for the second session.

On completion of the CATs the students filled in the questionnaire on their own. Follow-up telephone interviews were conducted during August 1996. The students were advised that the purpose of these interviews was for research on computer-assisted instruction and would not affect their student records in any way.

4.8 Summary

This chapter describes the ongoing formative evaluation of the tutorial “Random Variables” which took place during four micro-stages of the instructional design model:

- expert review
- prototype
- pilot test
- field test.

Computer-assisted instruction experts and subject matter experts evaluated the paper-based storyboards and the prototype computer-based product in a cycle of review, editing and further review.

The main instrument of measurement for the pilot test and the small-scale field test was an extensive questionnaire that had been carefully designed, so as to be totally

customised for the particular computer-assisted tutorial being evaluated, namely "Random Variables". The questionnaire addressed all the research questions and sub-questions of this study.

The pilot test was conducted in January 1996 with five members of staff from the Department of Statistics at Unisa. In March 1996 the small-scale field test was conducted in the computer laboratory at Unisa, with 25 students registered for STA101-H. The students were volunteers and were offered a financial contribution towards their travelling expenses. Being a distance education institution, it is difficult to obtain a random representative sample, due to the geographical dispersion of the students.

Follow-up telephone interviews were conducted with the same group of student volunteers during August 1996. The interview was also used as an opportunity to survey the computing facilities available to students, which will inform future multi-media developments.

The findings of the formative evaluation are discussed in Chapter 5.