A Case Study of the Use of Videoconferencing in a Speech and Language Therapy Setting

Elisabeth Mary Katsavras

A THESIS SUBMITTED FOR THE DEGREE OF MASTER OF LETTERS IN THE DEPARTMENT OF PSYCHOLOGY, FACULTY OF SOCIAL SCIENCE, UNIVERSITY OF GLASGOW, APRIL 2001.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Abstract

The subject of this thesis is the comprehensive evaluation of the telemedical TeachSpeech project which was a field study, designed by speech and language professionals. The TeachSpeech model of service delivery increases the therapeutic role of educational support assistants (ESAs) through the introduction of a videoconferencing link, used by a remote speech and language therapist (SLT) to support the ESAs in their work with children who have speech and language impairments in mainstream schools. This model of service delivery was contrasted with the traditional model of therapy where the SLT visits the schools to provide direct face-to-face therapy to the child. In contrast with the TeachSpeech model of therapy where the ESAs receive formal support across the videoconferencing link, the ESAs in the traditional model receive little formal support from the SLT.

The two models of therapy were evaluated through consideration of the project performance, the users' perceptions and the communication process of the videoconferencing meetings. The performance was measured using log sheets to gauge how the therapists spent their time, by a cost analysis and with a clinical effectiveness tool as applied to the nine children in the TeachSpeech group and the sixteen children in the traditional model of therapy. Stakeholders were canvassed using questionnaires and semi-structured interviews to appraise their perceptions of the models of therapy. Finally, the process of TeachSpeech model of therapy was assessed using structural and content analysis applied to the transcripts obtained from video-mediated meetings.

Results obtained with Enderby Outcome Measures showed that the two methods of delivery were equally clinically effective, thereby satisfying a critical requirement of the project. The technology was found to be unobtrusive, allowing the participants to adapt quickly to this innovative communication medium. The tech-

nology was found to support complex interaction such as the essential teaching-learning relationship that exists among the participants. In the small-scale pilot project considered, the TeachSpeech project was not cost effective. Under certain circumstances, extension to a wider population is believed to render the model cost effective. Results were mainly positive and suggested that consideration of a single evaluation criterion can be misleading. Despite being a small scale evaluation, the results were generally encouraging and suggest the potential benefit of this method of service delivery.

Contents

Abstract ii

Contents iv

List of Tables viii

List of Figures x

Acknowledgements xi

1 Introduction 1

1.1 Introduction to Thesis 1

1.1.1 Overview of Thesis 1

1.1.2 Summary of Results 1

1.1.3 Design of the Study: Strengths and Limitations 3

1.1.4 Outline of Chapter 5

1.2 Models of Therapy: and Their Efficacy 5

1.2.1 The TeachSpeech Project 5

1.2.2 Models of Therapy: Differing Intervention Agents 8

1.3 Telemedical Projects and Their Evaluation: Overview and Limitations 14

1.3.1 An Overview of Telemedicine 14

1.3.2 Telemedical Projects and Their Evaluations 15

1.4 Mediated Communication 22

1.4.1 Capturing Communicative Processes 23

1.4.2 Capturing Participant Performance 27

1.4.3 Capturing Perceptions of Participants 28

1.4.4 Summary 30

1.5 The TeachSpeech Evaluation 30

1.6 Overview of Thesis 32

2 Evaluating Performance **33**

2.1 Introduction 33

2.2 Overview of Chapter 33

2.3 Clinical Effectiveness 34

2.3.1 Method 34

2.3.2 Results 39

2.3.3 Discussion 39

2.4 Allocation of Therapists' Time 42

2.4.1 Method 43

2.4.2 Results 44

2.4.3 Discussion 45

2.5 Cost Analysis 47

2.5.1 Method 47

2.5.2 Results 48

2.5.3 Discussion 53

2.6 Conclusions 56

3 Evaluating Perceptions **57**

3.1 Introduction 57

3.2 Overview of Chapter 57

3.3 Method 58

3.3.1 Stakeholders 58

3.3.2 Questionnaires 58

3.3.3 Interviews 59

3.4 Results 59

3.4.1 TeachSpeech Group - Questionnaires 59

3.4.2 Comparison Group - Questionnaires 63

3.4.3 Interviews 65

3.4.4 Comparison Group - Interviews 66

3.4.5 Demonstration Programme Questionnaires 66

3.5 Discussion 67

3.5.1 TeachSpeech Group 67

3.5.2 Comparison Group 68
3.5.3 Demonstration Programme Questionnaires 69
3.6 Conclusions 69

4 Evaluating Process **70**
4.1 Introduction 70
4.2 Overview of Chapter 71
4.3 Coding Schemes 71
4.3.1 Interaction Process Analysis 72
4.3.2 Design Meeting Activity Coding 75
4.3.3 Conversational Games Analysis 78
4.3.4 Summary of Coding Schemes 80
4.4 Development of a Coding Scheme 81
4.5 The Current Coding Scheme 82
4.6 Method 87
4.6.1 Participants 88
4.6.2 Equipment 88
4.6.3 Transcripts 88
4.6.4 Reliability of the Coding Scheme 89
4.7 Results 89
4.7.1 Descriptions of the Dialogues Utilising Content Coding and Structural Analysis 89
4.7.2 The Use of the Visual Channel 96
4.8 Discussion 97
4.8.1 Do the Dialogues Change Over Time? 99
4.8.2 Explicit use of the visual channel 100
4.9 Conclusions 100

5 Conclusions **102**
5.1 Conclusions 102
5.2 Further Possible Exploration of the TeachSpeech Evaluation 105
5.3 Other Avenues 106

References **108**

Appendix A Specific Details of the Impairments of the Children in the TeachSpeech Group 120

Appendix B Specific Details of the Impairments of the Children in the Comparison Group 121

Appendix C Cost Analysis as a Pilot Project 122

Appendix D Cost Analysis as a Service 125

Appendix E Analysis of Questionnaires Given to TeachSpeech SLT 128

Appendix F Analysis of Questionnaires Given to the TeachSpeech ESAs 136

Appendix G Analysis of Questionnaires Given to the TeachSpeech Parents 145

Appendix H Analysis of Questionnaire Given to the Comparison SLTs 150

Appendix I Analysis of Questionnaires given to the Comparison Parents 156

Appendix J Professional Evaluation Questionnaire 160

List of Tables

1.1 A Comparison of Contact in the TeachSpeech and Traditional Models of Therapy 9

2.1 Recommended benchmarks for the interpretation of Kappa values . . 38
2.2 Median scores of first and final assessments on the Enderby Outcome Measures for both conditions 40
2.3 Results of EOMs at first and final assessment for the TeachSpeech group 40
2.4 Results of EOMs for first and last assessment of the Comparison Group 40
2.5 The average time in minutes set aside each month for each child, by the SLT 45
2.6 Resources and the costs for each model of therapy 49
2.7 Cost of Delivering the TeachSpeech and Traditional Model as a Service 50

3.1 TeachSpeech questionnaire completion rates 60
3.2 Comparison group questionnaire completion rates 64
3.3 Direct comparison of the satisfaction levels of parents, across both service delivery models, regarding the progress of their children . . . 64
3.4 Illustrates when and with whom the interviews were conducted . . . 65

4.1 A brief description of first level categories of Olson *et al.*'s (1992) design meeting coding scheme 76
4.2 Categories utilised in the current coding scheme 84
4.3 Percentage of turns allocated to each category averaged over the seventeen sessions 90
4.4 Percentage of turns allocated to each category in each session. 90

4.5 Structural analysis of dialogues, averaged over the seventeen sessions 91
4.6 Structural analysis of dialogues, illustrated for every dialogue. 92
4.7 Illustrates which ESAs contributed to which time zones and how much experience they had 94
4.8 Mean numbers of words, turns and words per turn over time 95
4.9 Mean number of turns and words split over time for each category of interest 95
4.10 Example of the use of the video channel - 1 96
4.11 Example of the use of the video channel - 2 97

List of Figures

1.1 Network of videoconferencing systems 8

2.1 The costs incurred as the distance the SLT has to travel increases. . 52
2.2 The costs incurred as the number of children increase. 52
2.3 The costs incurred as the life span of the equipment is increased. . . 54
2.4 The costs incurred if the travel time of the therapist, the number of children being treated and the life span of the equipment all increase. 54

4.1 The system of categories in Bales' Interaction Process Analysis . . . 73

Acknowledgements

I would firstly like to sincerely thank my supervisor, Anne Anderson, University of Glasgow, for the generous amounts of support and guidance she has given me.

There are also those who worked on the TeachSpeech project who indirectly provided guidance for my work. I would like to thank Karen Montgomery, I CAN, who delivered much needed enthusiasm and resources, and also Duncan Rose and Stephen Furner, BT Laboratories, who provided friendly and useful advice.

Finally, I would like to thank my family, friends and colleagues. I would like to especially thank my mother and father who gave me copious amounts of encouragement. I am extremely grateful to Todd Kleckner, who provided invaluable assistance and wisdom, and Rachel McEwan, who contributed much advice and direction.

ELISABETH MARY KATSAVRAS

Declaration

I declare that this thesis is my own work carried out under normal terms of supervision.

Chapter 1

Introduction

1.1 Introduction to Thesis

1.1.1 Overview of Thesis

This thesis describes the method by which a comprehensive evaluation tool was developed and applied to a telemedical project. The evaluation developed here was designed to address the limitations present in earlier research. In contrast to these earlier studies, the broad scope of the current evaluation seeks to overcome the narrow focus often criticised in previous evaluations. Thus, the evaluation tool developed includes analysing many aspects of the project and serves to provide a comprehensive summary of the relative merits of the programme.

This evaluation tool was applied to the TeachSpeech project. In this project distance support, using videoconferencing, was provided to workers within the speech and language therapy sector. Three evaluation criteria were used to compare the TeachSpeech model of therapy and the traditional model of delivery of speech and language therapy; the performance of the programme, the perceptions of those associated with the programme, and the potential impact on the process of communication of the videoconference technology.

1.1.2 Summary of Results

A major focus of the work described in this thesis involves a review of existing evaluation schemes. This not only serves to provide information about existing evaluation results and criteria, but also provides insight into effective ways to measure these criteria, and to judge in which scenarios these criteria are particularly relevant. By

considering these studies, and by taking into account the limitations of previous evaluations, a comprehensive evaluation method was constructed to assist in deriving conclusions concerning the overall merits of the TeachSpeech method of delivery.

The evaluation method compares a group of children receiving therapy delivered by conventional means and a comparable group of children which receives the TeachSpeech model of therapy. This involves measuring the clinical effectiveness of the video-mediated therapy compared to the traditional model of therapy, how the therapists apportioned their time, and the cost-effectiveness of the project. The TeachSpeech model of therapy was found to be as clinically effective as the traditional model of therapy. This result is of paramount importance to the future viability of the project, as a reduced clinical effectiveness would severely limit practical application of the project on a wider scale. It was found that, although there was no overall difference in the total amount of therapists' time utilised between the two models of therapy, the TeachSpeech therapists spent less time travelling and as a result were able to allocate more time to tasks directly related to the delivery of speech and language therapy. Finally, the economic performance of the project was evaluated by quantifying fixed, direct and indirect-variable costs of the project as compared to the traditional model of therapy. In the pilot project, the total cost of the TeachSpeech project exceeded the traditional model of therapy.

Closely related to the performance of the models of therapy were the perceptions of those people involved in the project. Speech and language therapists (SLTs) and educational assistants alike found that the technology was both convenient and easy to use. All those canvassed, either through interviews or questionnaires, expressed wide satisfaction with the project's ability to deliver speech and language therapy support. As these opinions included those of experts within the speech and language therapy field, this strengthens this finding.

Finally, a content coding scheme and a detailed structural analysis were used to examine the process of communication. These analyses suggested that users appear comfortable with the technology. They spent little time discussing the technology itself and the majority of time was spent discussing issues relevant to speech and language therapy. The analysis showed that a portion of the session time was used for social interaction, from which both teaching and learning can benefit. Finally, content coding of dialogues from the video-mediated sessions indicated the

need for a visual channel in specific circumstances such as being able to see the children's work.

1.1.3 Design of the Study: Strengths and Limitations

It is important to understand the context and limitations of the project to appropriately interpret the results of this study. The TeachSpeech project was designed by speech and language therapists with the aim of achieving the delivery of quality speech and language therapy. The therapists attempted to identify the best technological alternative to face-to-face interactions and thus high-quality videoconferencing was utilised. As the TeachSpeech project was designed before and not together with the evaluation project, the resulting design may be considered less than optimal for conducting a scientific investigation. In particular as a live case study it has characteristic strengths and weaknesses associated with a field-type study and these are discussed below.

The aim of the professionals who designed the project was to reduce the amount of time SLTs need to travel, while still providing quality speech and language therapy to the patients. The designers of the project elected to pilot the use of videoconferencing technology. They also decided that educational support assistants (ESAs) would have an enlarged role in the delivery of the therapy. Therefore, the resultant design of the TeachSpeech project allows only the evaluation of a model of therapy where videoconferencing technology is utilised and where the ESA has a more prominent role in the delivery of the speech and language therapy. For the purposes of a scientific study it would have been useful to investigate several further issues. Firstly, the addition of a no treatment group would have allowed an assessment of alternative models of therapy compared to no treatment, for instance, do children improve over time with no intervention. Secondly, using the videoconferencing link to provide direct therapy to children would have been useful. Thirdly, if other technologies such as audio-only communication could have been included their effectiveness compared to videoconferencing could have been judged. If audio-conferencing had proved effective, significant cost benefits could have been made. Finally, a larger sample size would strengthened any results obtained from the evaluation.

An advantage of the evaluation of the TeachSpeech project included being

able to capture naturalistic data, where the participants involved in the videoconferencing sessions were motivated by their professional needs to treat the trial very seriously. Clearly, the SLTs and ESAs involved in the project are interested in the outcome of their sessions and wish the outcome to be successful. The strength of naturalistic data lies in the ability to more reliably predict how users will respond to new technology in the context of their workplace. Furthermore, due to the lack of field studies in this area (Anderson *et al.*, 1997; Tang and Isaacs, 1993), the TeachSpeech project may be seen as particularly valuable, as it investigates a real scenario situated in the workplace over an extended time period. Tang and Isaacs (1993) suggest that video-mediated communication tends to have the most impact on interpersonal communication which takes place in a realistic setting. For these reasons it can be seen that the TeachSpeech project has several advantages over a laboratory-based study.

Although there are advantages in a field-based study there are also limitations. One such limitation in this clinically sensitive area, concerns not being able to manipulate the experimental variables of the study. Clearly, the SLTs' aim is to provide the best care possible in order to improve the clinical well-being of their patients. Likewise, the main priority of the patients and their care-givers is to see the successful treatment of the speech and language impairment. It seems logical to provide and test an alternative service which may be as effective as, or possibly more effective than, the model of therapy currently in place. The professionals who were responsible for the design of the TeachSpeech project decided it would be inappropriate to include a comparison group of children with speech or language difficulties who received no treatment. Therefore, this evaluation study compares the effectiveness of a model of therapy which utilises videoconferencing technology and an increased role for ESAs with the model of therapy currently in place.

The TeachSpeech method of delivery essentially combines two factors; the ESAs receive support and guidance from the SLT across a videoconferencing link, and the ESAs become the primary providers of support to the children with speech and language impairments. This has several repercussions for the design of the evaluation: unfortunately audio recordings of equivalent face-to-face sessions between the SLT and the ESA are very difficult to capture since in the comparison group this support is delivered in an informal manner and is not seen as critical as in

the TeachSpeech group where the ESAs are the main providers of therapy. The analysis of communicative process relies therefore on an investigation of only the TeachSpeech dialogues compared against benchmarks as set out in the literature.

1.1.4 Outline of Chapter

The remainder of this chapter discusses the background information relevant to the evaluation of the TeachSpeech project. Section 1.2 provides an overview of the TeachSpeech project and also other models of therapy where a third party delivers the speech and language therapy. Section 1.3 reviews the literature relevant to service-based projects, their evaluations, and the limitations of these evaluations including application to, and evaluation of, video-mediated projects. Section 1.4 reviews the effect of different technologies on the communication process with the aim of finding an evaluation tool that accurately measures the effect of mediation. Section 1.5 gives an outline of the conclusions reached from the literature reviewed and demonstrates how these conclusions influence the design of the TeachSpeech evaluation. Finally, Section 1.6 provides an overview of the remainder of the thesis.

1.2 Models of Therapy: and Their Efficacy

In order to understand the evaluation criteria to be implemented in Chapters 2, 3 and 4, the TeachSpeech project is outlined. There are two main components within the TeachSpeech project section, an overview of the traditional model of therapy and an overview of the TeachSpeech model of therapy. Following this an overview of models of therapy where an intervention agent other than a SLT delivers the therapy is presented. These overviews also serve to provide motivation for the evaluation of this project.

1.2.1 The TeachSpeech Project

The TeachSpeech project began in March, 1997 and lasted for three years. It was a collaboration between Invalid Children's Aid Nationwide (I CAN), the Department of Trade and Industry (DTI) and British Telecom (BT). This innovative project piloted the use of videoconferencing systems to provide speech and language support to schools within the Wiltshire area. The existing traditional therapy service is under pressure as more children are in need of speech and language therapy and there are

not enough SLTs to meet the demand. These factors are accentuated in Wiltshire and other rural areas since the SLT must travel large distances to and from schools.

The objectives of the project were:

• to supplement and extend existing I CAN outreach work to children with speech and language impairments by the use of videoconferencing technology, and

• to provide a clear demonstration to key UK decision makers of the benefits and opportunities offered by such multimedia services.

The following two sections outline in detail the two models of therapy which are to be evaluated.

The Traditional Model of Therapy

To fully understand the design of the present study it is helpful to examine the working environment of a SLT. The SLT would typically be employed by the National Health Service and, in relation to the TeachSpeech project, would use the traditional model of therapy delivery. SLTs work in wide ranging areas, including diagnosis, assessment and treatment of communication disorders associated with speech, language and swallowing problems. Their objective is to teach clients or patients to communicate more effectively within their environment. The treatment is carried out using a variety of techniques, such as intensive one-to-one or group therapy and, for children, is often carried out using play. The SLT initially determines what needs to be accomplished, and then works on areas of specific difficulty by developing a treatment plan based upon the individual's strengths and pre-existing skills. The treatment goals that are developed are often termed treatment outcomes. For children, speech and language therapy could involve developing language, improving articulation or phonology, strengthening oral motor skills, or improving fluency. For adults, speech and language therapy would focus on the same areas and, in addition, might also aim to improve oral motor muscle tone, motor planning skills, or refresh previously acquired skills. SLTs construct treatment plan outcomes and these are discussed with the patient and/or the family member in advance of the initiation of the programme. Treatment outcomes are continually modified based upon the progress the client demonstrates over time. Additionally, if a therapeutic approach is not entirely successful, regular monitoring of procedures and tasks used in therapy enables the patient and therapist to change the goals or methods to accomplish

new treatment outcomes. Professionally written narrative reports document the purposes and results of therapy and are an essential component of patient care.

Another aspect of a SLT's work pertains to their contact with various individuals within the social and medical services. Frequently, SLTs liase with such professionals as ear, nose and throat specialists, general practitioners, teachers, neurologists, psychologists, special educational needs co-ordinators (SENCOs), social workers and ESAs. SLTs often form part of a multi-disciplinary team.

As part of an emerging general educational trend, ESAs have been employed to help with expanding class sizes. There is no formal training for ESAs, although they often have had experience working with children. One of their many duties is to spend time working with special needs children. How time is allocated to this work is the responsibility of each SENCO. The SENCOs receive reports compiled by various professionals who have assessed the needs of the particular child and then they will consider how much individual time that child will receive from an ESA; typical per child allocations are fifteen minutes per day with the ESA. Within the traditional system, the SLT might visit a child every week and spend perhaps half an hour alone with that particular child in one-to-one therapy. In some cases, the SLT will consult with the ESA, but sufficient time may not always be available. In practice, the SLT rarely gets to see the child once a week, unless the case is considered severe. A more likely scenario would involve the SLT seeing the child once a month or perhaps even less often. This is because SLTs' caseloads are generally large and therefore the SLTs are under time pressure. Helge (1983) states that the recruitment and retention of SLTs has reached critical proportions. Other models of speech and language therapy have been developed which do not rely solely on the SLT delivering the treatment. An overview of one possible alternative will be given in the following section.

The TeachSpeech Model of Therapy

In the TeachSpeech model of therapy, the therapist visits the schools at least once a term to assess the child and liase with the teacher and ESA. In addition, these visits are supplemented by weekly videoconference meetings between the ESA and the SLT. These weekly sessions aim to report on the child's progress, make the appropriate adjustments to treatment outcomes and, if required, teach the ESA speech and language techniques to use. The aim of this evaluation was to assess

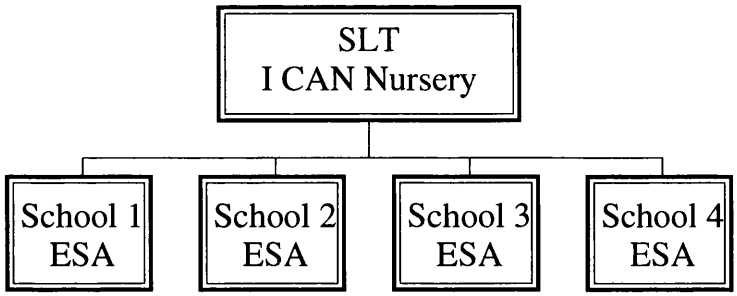


Figure 1.1: Network of videoconferencing systems

whether I CAN's model of therapy can be justified in a more permanent form.

As part of the TeachSpeech project, BT videoconferencing systems were installed in an I CAN speech and language nursery and four terminals were also installed in mainstream schools, as shown in Figure 1.1. A SLT was located at the I CAN Nursery and they videoconferenced with four mainstream schools once a week to provide support to ESAs, who were working with the children.

The children who participated in the project as part of the TeachSpeech group attended the I CAN nursery and they were then being integrated into one of the four mainstream schools which was taking part in the TeachSpeech project. The comparison children attended the same mainstream schools as the TeachSpeech children and they were also receiving speech and language therapy.

The TeachSpeech project is, therefore, an alternative service delivery model of speech and language therapy that utilises a technological solution. The emphasis of the therapist's role changes from providing direct therapy to a child, to supporting an ESA across a videoconferencing system. Table 1.1 illustrates this. To design an effective evaluation, it is necessary to investigate other evaluation studies which have utilised intervention agents other than the SLT and those studies that include broad evaluation strategies.

1.2.2 Models of Therapy: Differing Intervention Agents

The shortage of SLTs, the need to consider cost effectiveness and the potential to offer an accessible service to rural communities, have all been important incentives for piloting different service delivery models. Various models of speech and language therapy have been developed which utilise resources, such as education sector

Table 1.1: A Comparison of Contact in the TeachSpeech and Traditional Models of Therapy

Contact	TeachSpeech Model	Traditional Model
SLT and ESA	Video-mediated. Formal, usually weekly sessions.	Face-to-face. Informal, whenever there is time after visiting a child.
SLT and child	Face-to-face. Minimal only to appraise progress. Once a term.	Face-to-face. Depends on child's needs. Usually monthly.
ESA and child	Decided by SENCO. Depends on various factors including the child's language impairment.	

personnel, parents of children who are speech or language impaired and innovative teaching technologies. The investigation of the effectiveness of speech and language therapy, and its various service delivery models, however, remains in its infancy (Roulstone *et al.*, 1999). Nonetheless, an overview of the available studies, outlining how different service delivery models have been evaluated and the results obtained, is highly relevant to the present research. This section will firstly outline studies which have investigated whether an intervention agent other than a SLT is effective in delivering speech and language therapy. Secondly, an outline of studies that directly compare the support provided by the SLT to that provided by another intervention agent will be given. This section aims to appraise the type of evaluation methods used within this literature and explore the effectiveness of utilising another intervention agent in delivering speech and language therapy.

Early studies seemed to indicate that home-based therapy, where a parent delivered support, was ineffective (Cooper *et al.*, 1979; Kysela *et al.*, 1981). However, these studies were criticised on the grounds that no formal training in intervention techniques was provided for the parents (Snyder-Mclean and Mclean, 1987). More recent studies have substantiated the efficacy of parental input (Broen and Westman, 1990; McDade and Varnedoe, 1987; McConkey, 1985; Weller and Mahoney, 1983). Weller and Mahoney (1983) evaluated a project where the parents of Down's Syndrome children were trained to administer the Environmental Language Intervention Programme (MacDonald *et al.*, 1974). Parents received a 4 to 8 hour training session where they were instructed in the procedures they would use in the programme. The children, who were aged between 18 and 36 months, were all at the

one-word production stage and had no severe motor or sensory impairments. The clinical effectiveness of the programme was measured using a battery of tests. The results showed the children made significant progress on all cognitive and language measures and that their language quotients increased.

A study carried out by Broen and Westman (1990) looked at the effectiveness of parental input on children with phonological impairments. The results of the intervention provided by the parental input group was compared to two non-intervention conditions. These conditions were composed of a standard no intervention control group and a second group who were given parental intervention, this group was also monitored for six months prior to intervention to serve as the second no intervention condition. At the onset the children in the experimental group were aged between 43 and 60 months and their mean score on the Peabody Picture Vocabulary Test-Revised (PPVT-R) (Dunn, 1981) was 102.12. The control group were aged between 43 and 56 months and had a mean score on the PPVT-R of 101.9. All children were tested a total of three times, at 6 month intervals, using the Predictive Error Test (Westman and Broen, 1989). This test is a screening procedure designed to identify children with delayed phonological development. The results indicated that the experimental group, during the 6 months of intervention, advanced significantly more compared to both the control conditions.

The two studies outlined above illustrated how the clinical effectiveness of parental intervention, as compared to non-intervention, was assessed. Both studies used a pre- and post-testing system, while the number and type of tests differed across the two studies. Broen and Westman (1990) and Weller and Mahoney (1983) concluded that parental input could be effective when compared to a model where no therapy was given. The evaluation measures employed by the studies are, however, narrow and only address the clinical effectiveness of the models of therapy. Many questions are left unanswered, for example, whether therapy provided by a third party is as clinically effective as that provided by the SLT.

There seem to be relatively few studies that directly compare the effectiveness of different intervention agents in providing therapy. The studies which will be outlined below include a pioneering study by Cooper *et al.*, (1979) and more recent studies conducted by Ruscello *et al.*, (1993), Fey *et al.*, (1993) and Gibbard (1994).

Cooper *et al.* (1979) addressed the issue of comparing a home-based pro-

gramme where parents deliver the therapy, to a clinic-based programme where SLTs deliver the therapy. The authors specifically developed the Developmental Language Programme which was used to support the two groups outlined above. For a description of this programme see Cooper *et al.* (1978). Also, included in this study were a non-intervention control condition and a condition that received traditional therapy. A structured therapeutic programme was implemented in a clinical setting where SLTs provided group therapy for two hours per day, five days a week. The same programme was implemented in a home setting, where the work occurred daily with the child on a parent-child basis, and a SLT visited the home once every six weeks. The effectiveness of the models of therapy was assessed using a clinical measure, the Reynell Developmental Language Scale (Cooper *et al.*, 1979). The authors concluded that the clinic-based programme was most successful, followed by the home-based model. The traditional and non-intervention conditions, in terms of clinical effectiveness, fell short of the home-based condition. The non-intervention condition was assessed as the poorest. However, this study has been criticised (Law, 1997; Fey *et al.*, 1993). Law (1997) suggested that the amount of time in therapy offered to the clinic-based condition was considerably longer than that in the other conditions and Fey *et al.* (1993) suggested that the procedures used to assess subjects and the measures used to evaluate the clinical effectiveness were problematic.

More recently, Ruscello *et al.* (1993) conducted a study which directly compared two different models of service delivery. There were two conditions in the study. Children in one condition received therapy which was provided exclusively by a SLT and children in the second condition received a combination of a therapist administering therapy, parental involvement and a computer-based learning tool, called Speech Viewer. The children who took part in the study had a mean age of 61 months and a score in the 15th percentile or lower for their age on the Khan-Lewis Phonological Analysis (1986). The children in both groups received a total of 16 one-hour lessons over a period of eight weeks. The children in the control condition spent all 16 sessions with a SLT, while the children in the experimental condition spent eight sessions with the SLT and eight sessions with their parents and the Speech Viewer. To measure the effectiveness of the models of therapy pre- and a post-clinical tests which involved audio recording samples of the child's speech were administered. The parents completed a questionnaire which assessed their percep-

tions of the system and its utilisation. Results showed that both groups made gains and that these gains were similar in magnitude. The questionnaire data indicated that Speech Viewer was not difficult to use, the parents felt that the system helped their children and they would be willing to use the system again. The evaluation of this study is more thorough than those previously mentioned as it considers both the perceptions of the parents and the clinical effectiveness of the respective models of therapy. However, the limitations of the study seem to lie in a lack of a more practical evaluation tool, that of cost analysis.

Fey *et al.* (1993) undertook an evaluation of two different service delivery models of therapy which aimed to provide grammar facilitation and compared them to a non-intervention control group. This evaluation included a cost analysis. In one model therapy was administered by a SLT, while the participants' parents administered the therapy in the other model. Over a four-and-a-half month period, children who were aged between 44 and 70 months and who had a marked delay in grammatical development, were treated in one of the two conditions. The children's impairment was assessed thoroughly using a battery of tests and the parents who were to provide support were trained in the appropriate therapeutic area. To evaluate the clinical effectiveness of the respective models of therapy, thirty-minute language samples were elicited from the children. These samples were transcribed and analysed using the Developmental Sentence Scoring Module of Computerised Profiling Version 6.2 (Long and Fey, 1989). The first statistical analysis that was conducted suggested that there were substantial effects for the treatment groups as compared to the delayed-treatment control group, but there were no differences between the parent-based treatment and the SLT-based treatment. However, further analysis indicated that the clinical effectiveness of treatment provided by the parents was not as consistent as that of treatment provided by the SLT. The authors suggested confounding factors might have been the training methods used in the parental programme. As mentioned previously, a cost analysis quantifying the therapists' time was included in this study. Records were kept by the project SLT concerning the amount of time the clinicians spent facilitating the respective models of therapy. Results showed that the SLTs treatment group required 240 hours of therapists' time for a group of six children, while the parental treatment group only required 126 hours of therapists' time, indicating a saving in service provision costs.

This study did not systematically assess the perceptions of the users involved in the different models of therapy.

Gibbard (1994) also investigated cost effectiveness and addressed what is termed the Hawthorne effect. The Hawthorne effect occurs when therapeutic effects, while evident, are thought to be the result of contact with a sympathetic individual rather than the result of any intervention technique (Law, 1997). Gibbard's (1994) study assessed four conditions: a delayed-intervention condition, a condition which received traditional therapy, a condition which involved patients receiving therapy from parents who had been trained specifically in speech and language therapy techniques and finally a condition which received support from parents who had been given a general non-specific developmental training. This study was conducted over a six-month period with children who were aged between 27 and 39 months and who were considered to have a delay in their expressive language skills. Clinical effectiveness was measured by the Reynell Developmental Language Scales (Reynell, 1983), the Renfrew Action Picture Test (Renfrew, 1986), the Derbyshire Language Scheme Picture Test (Knowles and Masidlover, 1979) and a language sample. The findings showed that, when the parents were specifically trained in speech and language techniques, they were as clinically effective as the speech and language therapists. As a result of her evaluation, Gibbard (1994) concluded that utilising the parents as intervention agents was a cost effective method of delivering speech and language therapy. She suggested that SLTs could usefully adopt a more consultative role, since the subsequent effects of this could allow for an increase in the number of children being treated.

The studies outlined above seemed to suggest that intervention by an agent other than the SLT can be effective. Results directly comparing the clinical effectiveness of a SLT to that of another intervention agent are less conclusive. Although one early study (Cooper *et al.*, 1979) indicated that parental intervention is not as effective as therapists intervention, for the most part studies have contended that parental intervention can be as effective as therapists (Gibbard, 1994; Ruscello, *et al.* 1993; Fey *et al.*, 1993). The studies outlined have tended to concentrate on measures of clinical effectiveness and have not used broader evaluation strategies which would provide a more detailed picture of the models of therapy. It would seem that a wider literature needs to be examined in order to assess a wide range of evaluation

methods and results.

1.3 Telemedical Projects and Their Evaluation: Overview and Limitations

In this section, both historical and current telemedical programmes will be examined with an aim to consider the results obtained and to gain insight into the way in which such programmes are effectively evaluated. This literature is relevant since, like the TeachSpeech method of service delivery, it utilises technology to deliver medical opinion.

1.3.1 An Overview of Telemedicine

Field (1996) gives an example of forward thinking in 1924, where the front of an innovative magazine, Radio News, has an illustration depicting a radio doctor. The patient and doctor are participating in what would now be termed a videoconference. At the time of publication, radio was only just reaching American homes and experiments transmitting pictures had not yet occurred. While this was not realised for over forty years, the delivery of health care with a health expert and patient located at different sites has since occurred. The two main terms that capture this phenomenon are those of telemedicine (TM) and telehealth.

The literature does not necessarily agree on the definition of TM. Stamm and Perednia (1998), Garshnek *et al.* (1997) and Field (1996) use the U.S. Department of Commerce (1997) definition, which can be summarised as the use of electronic communications to deliver health care services at a distance. Using this definition, telephone consultations between a patient and physician or the use of the radio to link emergency personnel with a local medical centre would be included. Other definitions of TM exclude these situations; McLaren and Ball (1995) and Preston *et al.* (1992) use the definition "Telecommunication that connects a patient and a health care provider through live two-way audio, two-way video transmission across distances and that permits effective diagnosis, treatments and other health activities". Even though the many authors use different definitions, their papers tend to focus on studies that use technology other than the telephone and radio.

Telehealth includes TM and other health-care related activities, and refers to

the use of electronic communications and information technology to provide access to assessment, diagnosis, intervention, consultation, education and other information across distance (Nickelson, 1998). For the purposes of this work, it is most useful to focus on studies that use video technology in a telemedical setting, as telehealth encompasses health-service components which are not relevant to the current study.

1.3.2 Telemedical Projects and Their Evaluations

Whitten and Collins (1997) give a detailed review of how TM evolved. They believe that TM has had two births, the first being in the late 1950s and the second in the early 1990s. In the following sections, both of these births, namely the historical and current programmes, will be examined to understand the limitations of the evaluations of these programmes.

Historical Programmes

The first TM project is generally cited (Whitten and Collins, 1997; Field, 1996; Conrath *et al.*, 1983; Wittson and Benschoter, 1972; Benschoter *et al.*, 1967) as a project which was established in 1959 at the University of Nebraska. Wittson and his colleagues used two-way interactive video communication to transmit neurological examinations and other information to medical students who were situated at a remote site. In 1964, they used a similar link to provide case consultations and speech therapy diagnosis of difficult psychiatric cases. The link extended from the Nebraska Psychiatric Institute to Norfolk State Hospital that was situated 112 miles away. These researchers were able to show that the consultations over a long distance were no less effective than when the physician was physically present, although the measures used were not stipulated.

Another project which originated in the late 1950s is Space Technology Applied to Rural Papago Advanced Health Care (STARPAHC). This was a joint venture between National Aeronautics and Space Applications (NASA) and the U.S. Public Health Service, and aimed to provide medical services to Papago Indian communities. Justice and Decker (1979) compared the telemedical care programme to that being provided by clinics staffed by physicians within the same health system. They assessed the reliability of the equipment, patient acceptance and project costs and they found no significant differences in the quality of health care provided.

In 1971, the use of telemedical systems was surveyed by the National Academy of Engineering. These systems involved physicians assisting nurses who were providing home care and supervising non-physicians supporting ambulatory care clinics. The former application was used by the Mount Sinai School of Medicine in New York which, in 1972, established a black and white cable television link which supported nurses who were providing paediatric care in a Hispanic area of the city (Muller *et al.*, 1977). Also in the 1970s, the U.S. Public Health Service and the Department of Defence funded a series of teleradiology projects that involved transmitting radiologic images. These projects led to the development and implementation of civil and military teleradiology (Mun *et al.*, 1989). Details of these projects' evaluations were sparse.

According to Perednia and Allen's (1995) review of TM studies, only one project has survived since 1986. This project, established by the Memorial University of Newfoundland, began in 1977 as a three-month pilot project to test the usefulness of a two-way audio and one-way video system. After a three-month trial, the authors concluded that television was useful but found that the project team could exchange educational material more efficiently and less expensively by phone, videotape and printed materials (House, 1993). These findings are consistent with a comprehensive study conducted by Dunn *et al.* (1977), which compared the diagnosis of over a thousand patients by a physician. The physicians conducted their consultations by using the telephone, using still or motion frame black and white television, or colour television. No significant differences between the modes of communication were found for diagnostic accuracy, time of diagnostic interview, tests requested, referral rates or patients' attitudes. They concluded that the relatively inexpensive hands free telephone was as good a diagnostic instrument as the colour television system.

As can be seen, the initial surge of TM studies started in the late 1950s and continued until the mid 1970s. Preston *et al.* (1992) suggest that the main reasons that the TM phenomena did not spread was due to high costs associated with telecommunication technologies and due to the difficult challenges in organisation and management. The aforementioned TM projects lie in many different fields of medicine, yet it is found that the impetus behind many of these projects occurred where geographical factors make travel difficult and therefore expensive. Although

evaluation of these and other projects is scarce, the available studies suggest TM programmes are equally effective as traditional delivery methods. Both the patient and provider are generally satisfied with the outcome, although some call into question how technologically advanced systems need to be to support the patient's or the provider's needs.

Current Programmes

The second birth of TM projects occurred in the early 1990s and is generally attributed to recent technological advances such as fibre optics, integrated system digital networks (ISDN) and compressed video (Garshnek *et al.*, 1997). Such technologies provided audio synchronisation, improved video image quality and enhanced usability, while simultaneously reducing systems' cost. The closure of many rural health-care facilities has resulted in increased travel for medical purposes. In 1993, U.S. citizens logged over one trillion miles of travel for medical purposes (Wade, 1994).

It is difficult to track the current rate of expansion of telemedical projects. Field (1996) stipulates that since the area of TM is expanding at a rapid rate, this renders it difficult to obtain a complete inventory of current projects. Not only is the field expanding, but projects are funded through radically different bodies such as the government, the private sector and commercial organisations; the latter two are particularly difficult to track. Allen and Perednia (1996) suggest that the number of TM projects in the U.S. had reached 50 in 1995 and since then has doubled each year.

Projects involving TM can be divided firstly by medical speciality and secondly by the context of use. The medical field is diverse, and telemedical projects include teleradiology, telepathology, telepsychiatry, teledermatology and telesurgery. The two most common, teleradiology and telepathology, will be outlined below. The context in which TM is used includes TM projects located in prisons, in the military, and in space. These applications will be examined due to their similarity to the TeachSpeech project in that they use videoconferencing technology. As stated previously, the whole impetus behind TM is to cut costs by reducing travel time and to provide services to areas that are difficult to access. The examples of using TM in the military and in space present unique scenarios in which there are few alternatives

and distances are great. The effectiveness of these systems are potentially critical to the health and well-being of the users.

Teleradiology

Teleradiology, which involves radiologic image transmission, has been described as the compression and transmission of images and patient data and subsequent reconstruction of the data for interpretation at a remote site (Forsbery, 1995). Teleradiology is currently thought to be the most common application of TM as 67% of all telemedical consultations were attributed to teleradiology (Hoffert, 1997). There are several reasons given to explain the fact that teleradiology has become the most common use of TM. Historically, teleradiology relied on an established network of courier services to obtain second opinions, and after the advent of electronic transmission, this became to be seen as costly and inefficient. Secondly, in the U.S. payment of teleradiology consultants through medical insurance is available, while other forms of TM are not subsidised. Finally, radiology has a history of using digital imaging techniques, such as magnetic resonance imaging (MRI). Thus, a logical next step in this field is to send images digitally. Franken (1996) describes a programme at the University of Iowa where an already established film-based radiology programme progressed to a digitally based programme.

Several teleradiology studies have been evaluated. Fajardo *et al.* (1990) compared diagnosis of conventional mammograms with teleradiologically transmitted mammograms. Four specialists examined the images for skin and nipple abnormalities, microcalcifications and masses. While the results showed that the specialists were significantly better at reading the conventional mammograms for skin and nipple abnormalities, no differences were found when the specialists were detecting masses. It was suggested that if the image quality was improved, the detection of skin and nipple abnormalities would also improve. Likewise, Scott *et al.* (1993) evaluated a study which aimed to detect subtle orthopaedic fractures. Images were digitised and eight interpreters read the images. Overall, a significant difference was found as interpreters were consistently better at rating the original, analogue images. As such, the authors concluded that the system was not satisfactory for their purposes. Finally, Slovis and Guzzardo-Dobson (1991) studied a programme which used teleradiology to examine chest and abdomen abnormalities of new-born infants. In this case, images were read via video or via radiographic film by two

readers. Agreement between the two conditions was 98%. The authors concluded that the use of the system in this way was appropriate.

Telepathology

Telepathology is another main area where TM is used. The term telepathology encompasses the exchange of pathology images through telecommunication for the purposes of diagnosis, consultation, research and education. The use of telepathology is of great importance in the management of patients since it allows fast diagnosis and consultations between specialist pathologists located in every part of the world (Rosa, 2000). The first main method of telepathology is static telepathology, which captures and digitises an image as selected by a pathologist. The image is then transmitted by electronic means to a telepathologist. The second main method is dynamic telepathology, which consists of a live electronic connection between the sender and the receiver. As an example, a microscope capable of sending images may be equipped with a telerobotic system that is remotely operated by the telepathologist making the diagnosis.

Weinstein *et al.* (1989) published data on a telepathology programme which examined breast tissue. Six pathologists looked at 115 specimens by both light and video microscopy. Overall, the pathologists performed equally well although there were individual differences. The study concluded by cautioning others not to extrapolate these results to projects using different equipment or looking at different types of tissue. Another study of breast tissue specimens using videophone provided support for telepathology (Linder and Masada, 1989) and although only one pathologist read the specimens, they did so with 95% accuracy. A review of telepathology programmes offered by Grisby and Kaehny (1993) concluded by stating that telepathology is feasible, but evaluation studies so far indicate that the technology and image quality must be of a high standard.

Military

The military has initiated some of the most ambitious programmes to date (Garshnek *et al.*, 1997). Recently a network was established by the U.S. Department of Defence (DOD) for a telemedical system which serves U.S. troops based in Bosnia and other countries. The telemedical system is currently known as Operation Primetime III, and was implemented to aid army physicians. Communication between the

army physicians and a federal hospital was achieved through real-time audio and visual technology. The system uses a satellite which allows for direct broadcast quality for the purposes of consultation and diagnosis. Operation Primetime was first established in 1993 to support medical units in Macedonia and Croatia, and in 1995 was upgraded and renamed Primetime II with a 30-fold increase in communication bandwidth, which significantly increased the quality of transmission data. Primetime III now provides an integrated world-wide system of TM for the American military. The AKAMAI TM programme, used for both military and civilian personnel, was established by the military in 1993 in Hawaii (Delapain *et al.*, 1993). The TM sessions are used for diagnosis and consultation in a state with particularly high health-care costs.

Since the DOD uses TM extensively, they have been working to develop a coherent evaluation strategy. Walters (1996) examined 171 TM consultations received from deployments in Somalia, Macedonia, Croatia and Haiti between February, 1993 and March, 1995, where she looked at the purpose of the consultations as rated by experts. The experts suggested that in 30% of cases the consultation significantly changed the diagnosis and in 32% of cases the consultation changed the course of treatment. The experts' review described the consultations as essential or stated that they prevented evacuation in about 10% of the cases. The evaluation study also highlighted issues pertinent to other TM projects. Walters found that it was difficult to sustain the TM consultations over time for a variety of reasons: trained participants left, the technology felt awkward to use and the physicians at the remote sites became more informed during the initial consultations and as a result felt they did not need to consult for subsequent patients.

Space

Astrotelemedicine or TM in space is also a well-established concept. Originally, medical monitoring would include monitoring heart rate, oxygen consumption levels, heat production, carbon dioxide levels and other variables. These were monitored by a NASA medical team. Currently NASA has a permanent operational international space TM programme and both the Space Shuttle and Mir programmes have TM capabilities. Medical conferences can be held between the crew surgeon, located on Earth, and crew members during extra-vehicular activity. This system has the capacity to be an important lifeline, giving access to medical expertise and crucial

instruction on the new International Space Station (Garshnek *et al.*, 1997). Since there are no feasible comparisons to telemedicine in space, and to date no serious medical emergencies, evaluation has not taken place.

Prisons

A third instance of videoconferencing-based TM in restricted areas is in prison environments, where there is a great potential to reduce costs. Reasons for commencing telemedical projects in prisons include the expense required to transport patients, since guards and an appropriate vehicle are needed, or the expense required bringing specialists in due to adverse working conditions. Programmes are currently being run in Texas, Oklahoma, Colorado and North Carolina. East Carolina University (ECU) provides a TM service to a maximum security central prison. At the initial stages of the programme, two physicians worked at a facility 100 miles from the prison; the network that linked them to the prison was established in 1989 and was a state-wide distance learning network (Tichenor *et al.*, 1996). Currently, the programme includes 15 ECU physicians. A financial audit by the North Carolina Department of Corrections in March, 1994 found evidence of cost savings. Allen and Stein (2000) reported savings of $225 per month for each of 95 inmates in an Oklahoma prison. Evaluation studies here seem to concentrate on cost effectiveness issues and in general, there is a lack of data concerning clinical effectiveness and the participants' perceptions of the programme.

Evaluation critiques

There have been criticisms of how TM projects have been evaluated (Hoffert, 1997; Institute of Medicine Committee, 1997; McLauren and Ball 1995). The Institute of Medicine (IOM) Committee report (1997) suggests that progress within TM is hampered by the lack of reliable comparison studies. It believes that a rigorous standard of review should be applied to TM technology to determine its clinical and cost effectiveness. The committee suggests that systematic comparison between TM and its alternatives should be made, taking into account caregiver and patient satisfaction, health outcomes, amount of access, costs, and the satisfaction of clinicians. It also advises that, since TM may have a potential impact on staffing levels and responsibilities, many clinicians may see it as an economic threat, thus making evaluation more problematic. Hoffert (1997) concurs, emphasising the fact that despite

$646 million worth of TM projects, there are relatively few evaluation studies that look at economic issues. McLaren and Ball (1995) call for more rigorous evaluations and believe that the few evaluation studies seen as rigorous show results that suggest TM programmes are not cost effective, and that state of the art technology is not required. As an example of this, Moore *et al.* (1975) compared two media for remote diagnosis, namely the telephone and the television. They found that the television was no more effective than the telephone, although the television consultations were significantly longer and more expensive. This is consistent with Dunn *et al.* (1977), whose results compared the diagnosis of patients by a physician either face-to-face, or communicating via various different media. Very few differences were found.

From this, the general consensus is that sophisticated equipment is not required for many telemedical projects. However, it should be noted that these conclusions were drawn by mainly examining the clinical effectiveness of the programme. While this is one way in which insight might be gained into the relative merits of the programme, it is clearly not the only one. Monk *et al.* (1996) suggests taking a multi-dimensional approach when evaluating the effects of mediated communication. It seems clear that whether or not the communication medium has an impact on interaction between two parties is important. However, this has not been the focus of many of the previous studies in the field of telemedicine. In the TeachSpeech project communication between the SLTs and ESAs is critical. From other kinds of evaluation of technology, it has been shown that the process of communication has been affected by the introduction of technology. Based on this, the following sections examine how video-mediated communication is evaluated and whether differing communication channels affect the process of communication. This examination provides a review of the tools and techniques used to evaluate various technologies and also serves to illustrate the differences, one expects to find when communication is mediated using videoconferencing technology.

1.4 Mediated Communication

As stated above, telemedical projects tend to focus on performance measures and neglect other factors which might be affected by the introduction of new technologies. Telemedical studies have suggested that the visual channel has been found to provide little advantage as compared to services which utilise audio-only com-

munication. There is however an entire field of literature which not only examines the affect on performance of communication medium but also investigates the affect on communication. In this field the affects of mediation are assessed in three main ways: by detailing measures of task performance, by capturing the process of communication and by recording the perceptions of users (Anderson *et al.* 1999; Olson *et al.* 1997; Monk *et al.* 1996; Sellen, 1995). Clearly, each of these factors are relevant and important in considering the success of the TeachSpeech project. Thus, in the following, different techniques will be reviewed which have been applied to the analysis of mediated communication in order to assess their applicability to the TeachSpeech project.

1.4.1 Capturing Communicative Processes

To examine whether or not mediation has an impact on communication, we need to first be able to analyse communication in a variety of ways in order to fully understand the effects of mediation. From this, the potential usefulness of structural analysis and content coding with regard to the TeachSpeech project will be established. The first technique reviewed relies on a structural analysis of the dialogue while the second relies on content coding of dialogues.

Structural Analysis

Capturing the process of communication generally involves analysing transcriptions of audio recorded dialogues. Analysing transcripts can be done in various ways, including investigating the structure of dialogues. Sellen (1995) describes how to characterise communication using its surface structure. The surface structure is composed of three main components which allows one to analyse dialogues. The first component involves the total number of words spoken in a dialogue. The second component is termed a turn, where during any given turn one participant remains the principal speaker; attempts by another participant to become principal speaker which fail are termed interruptions. The third component is the mean number of words spoken per turn. How these measures might provide information about the communication process is best illustrated through example, as illustrated below.

Firstly, it has been shown by structural analysis that communications technologies which produce audio delay affect communication and, as a result, perfor-

mance. Tang and Isaacs' (1993) observed that audio delay can lead to difficulties in turn-taking, while similarly O'Conaill *et al.* (1993) found longer, fewer turns result from a delayed audio signal. In Anderson *et al.*'s (1997) study of the effect of transmission delay, they too found a significant disruption in turn-taking, and correlated this observation to the drastically reduced performance scores they measure. This is in agreement with O'Conaill *et al.*'s (1993) suggestion that the frequency of dialogue interruptions can be related to differences in task outcome. Thus, as a tool, structural analysis has been used to demonstrate the impact of one form of communication technology and its characteristics, namely audio delay, on communication and task performance.

Secondly, examples in the literature tie together observations of total dialogue length as a measure of communicative efficiency; Anderson *et al.* (1997) illustrate this through their assessments of a videoconferencing system which supported direct eye contact between users in which the total number of words required to complete a task greatly exceeds those of the face-to-face condition. Furthermore, Monk *et al.* (1996) hypothesise that the ease of communication can be measured through the number of words in a dialogue, while Sellen (1995) stipulates that communicative ease is best characterised through a low number of words per turn, as frequently observed in face-to-face interactions (O'Conaill *et al.*, 1993; Cohen, 1982). Again, structural analysis has been used to investigate the effect of mediation upon communication.

Finally, the work of Tang and Isaacs (1993) suggests that video-mediated dialogues are less interactive. The researchers observed less frequent changes of speaker turns in video conferencing compared to face-to-face meetings, which results in longer turns and less back-channelling. They inferred from this that, compared to face-to-face meetings, VMC suppressed complex, subtle, and difficult interactions where the participants felt reluctant to express their opinions and avoided working through disagreement.

In summary, these studies show that communication varies depending on whether it is mediated or not, and by the type of technology it is mediated by. These examples also illustrate some of the potential value and types of inferences that can result from the use of structural analysis within the VMC domain. In particular, and especially relevant to the TeachSpeech evaluation, structural analysis seems ideally

suited to determine the degree to which the technology is invisible to users and whether or not the technology impacts upon the interactivity of the communication. With this aim in mind, we seek to quantify the ease of communication as measured through words per turn and overlapping speech measures.

Content Coding

Another technique which measures the process of communication and can provide informative data is by coding the dialogues, to investigate the content or function of utterances. Coding dialogues can be a useful research tool as this type of analysis can reveal aspects of the interaction of which participants are unaware and it can show how the communication link is used during videoconferences. Content coding can indicate whether the technology is interfering with the interaction, for example, it may show that participants are talking about the technology 80% of the time and talking about the task for only 20% of the time. In this example the technology is impacting heavily on the communication. Content coding can also reveal how information is shared between participants. In a similar fashion to structural analysis, the data that are gathered are naturalistic and there is little danger that the measurement tool has affected the data. In the following, we describe the Conversational Games Analysis (CGA) coding scheme developed by Kowtko *et al.* (1991) and overview results obtained with this and other examples of coding schemes as applied to VMC dialogues.

It has generally been found that content coding differentiates between audio-only, video-mediated and face-to-face communication in the number of instances in which either the speaker or listener seeks to confirm understanding. In the context of CGA, and as compared to face-to-face communication, this translates into a high number of check and align games for audio-only communication, and frequent check games for VMC (Anderson *et al.*, 1997; Doherty-Sneddon *et al.*, 1997). Such investigations have been used to illuminate the reasons behind longer conversations in VMC as compared to face-to-face communication (Anderson *et al.*, 1997). In this sense, content coding has been used to complement the information afforded via structural analysis.

CGA is a type of coding scheme which primarily investigates how information is shared between participants. Audio recordings of the dialogues are transcribed,

and the content of these transcriptions is apportioned by coders between different categories, each of which is termed a game within the context of CGA. Within CGA, there are six different types of games, which are found both necessary and sufficient to encompass all the different possible types of dialogue that take place within a conversation. They are as follows: *instruct*, which describes a request; *check*, during which a listener checks for understanding of a previous message; *query-y-n*, which represents a yes-no question; *query-w*, which describes an open answer question; *explain*, for which information is offered freely; and *align*, where the speaker confirms the understanding of the listener. Based on these games, the coder is able to determine how much of the conversation time is spent within each category.

Other examples which highlight possible uses for content coding involve investigations of detrimental performance scores in an audio-only condition for a collaborative problem-solving task (Veinott *et al.*, 1999). Through addition of a video-mediated condition and subsequent coding, it was found that the performance loss could be attributed to less instruction and checking of participants mutual understanding in the audio-only condition. The final instance of content coding described here, and perhaps the most relevant to the TeachSpeech project, lies in Olson and Olson's (1992) coding scheme, in which they introduce new categories termed *technology confusion* and *technology management.* Technology confusion describes times during which confusion concerning the technology occurred, and technology management comments on the distribution of information on the computer screen. Although no quantitative data is presented regarding the amount of time that was spent discussing technological issues, this innovative concept could be used to determine to what extent the technology affects the desired interaction.

In the context of the TeachSpeech project, content coding could be valuable in assessing how much the technology distracts the SLT and ESA during their videoconferencing sessions. Furthermore, this type of analysis might also be able to provide a form of convergent validity, either confirming or not another finding. For example, users might report that they felt the technology did not impact upon their communication and this finding could be strengthened by content coding which found infrequent talk concerning the technology. It would seem that content coding is a useful addition and that it can capture a unique aspect of the communication.

1.4.2 Capturing Participant Performance

In this section we discuss techniques relevant to measuring task performance and the results obtained with these techniques. Clearly it is very important to realise when performance is compromised as a result of the communication channel, and in what follows we aim to review existing techniques that are potentially applicable to measuring performance scores for the TeachSpeech project.

From the literature surveyed, it is apparent that there are many criteria used to evaluate the performance of participants owing to the wide variety of tasks assigned; performance scores related to these tasks are often referred to as task outcomes. Task outcome provides information about whether the communication was effective in allowing task completion and to what standard the task was completed. Task outcome measures can include the time required to complete the task or a grade representing the quality of the task performed. Virtually independent of the criteria used, performance data were generally stable among the differing mediation channels for co-operative problem solving tasks (Anderson *et al.*, 1999; Anderson *et al.*, 1997; Chapanis, 1988). For example, while Anderson *et al.* (1997) assessed performance using different criteria depending on the task involved, they found that different communication media did not affect task performance; in these studies, performance criteria were generally comprised of accuracy scores and the overall thoroughness of the completed task. Likewise, Chapanis (1988) measured performance by recording the time for task completion and assessed the quality of the final outcome of the task. Their results also showed no difference between the audio and the face-to-face conditions.

However, there are exceptions to the trend established by the previously mentioned works. Specifically, Olson *et al.* (1997) found a performance score difference between an audio-only condition and a face-to-face condition in the quality of a completed design task, while VMC did not differ significantly from either of these two conditions. A second example which illustrates results in different performance scores for differing communication media is when tasks involve conflict. Tasks which involve conflict tend to exhibit performance differences (Clark and Brennan, 1991; Rutter, 1987; Williams, 1977), but are not addressed in this review owing to the stark differences that exist between these studies and the nature of the current, co-operative work. Unlike the obvious benefits afforded by an analysis of the process of

communication, performance analyses do not seem readily applicable to the TeachSpeech project. In the first instance, we are mainly interested in quantifying project performance through clinical effectiveness, and thus most of these performance criteria are not applicable. Secondly, the main aim of the project is to provide quality support for distance speech and language therapy, and thus in this particular scenario it is not feasible to manipulate the therapy environment as the quality of care may be compromised; this excludes the possibility of comparing performance measures for differing communication media. Clearly, studies exist which randomise treatment and they often include a condition in which participants receive no treatment. However, the intention of the TeachSpeech project was not to investigate whether or not a particular type of treatment technique was effective compared to a no-treatment condition. Rather the goal was to conduct a study whose design was chosen by professional who were concerned about maintaining high quality care throughout the trial. However, while not the primary concern, the therapy time required to achieve a particular clinical standard may be a useful measure to take into account any deficiencies resulting from different forms of mediation during the clinical sessions.

Clearly, obtaining performance measures constitutes a critical component of any evaluation study. Unfortunately, the effectiveness of the TeachSpeech videoconferencing meetings can not be directly assessed since no comparison of an audio-only or face-to-face meetings were available. However, long term global measures do exist. For example, performance can be assessed through the clinical outcomes of the children or the economical performance of the models of therapy. From the literature reviewed above it would seem that performance measures tend to remain stable across communication medium except in specific circumstances which do not apply to the TeachSpeech project.

1.4.3 Capturing Perceptions of Participants

Another measure that can be used to assess the success of communicative interaction is the perceptions of users. Questionnaires and interviews, are often used to gather subjective opinions regarding the participants' views, feelings about the nature of their interaction and any perceived influences of the communication medium.

In general in research in this field, questionnaire and interview data indicate

that face-to-face interaction tends to be the favoured method of communication while users also consistently report the subjective benefits of VMC as compared to audio-only communication (Olson *et al.*, 1997; Sellen, 1995; Tang and Isaacs, 1993). This is an important finding since it places VMC ahead of audio-only communication when choosing a mode of mediation, although as stated previously, performance tends to remain stable and independent of the form of this mediation. The results of Tang and Isaacs (1993) extend this result, and confirm that users believe the addition of the visual channel makes savings both in terms of travel and time, two issues particularly important to a project concerned with distance support. Thus, the participants perceived a benefit in utilising videoconferencing, this is an important finding when considering future uptake of the technology. In the questionnaire data of Daly-Jones *et al.* (1998), the preference of VMC to audio-only communication was found and, in some cases, VMC was rated the same as face-to-face communication. From this, the researchers concluded that VMC required less effort than audio-only communication. They felt that although participants were performing as well in audio-only situations, this was at the expense of greater communicative effort. Other areas in which researchers report that VMC can offer benefits compared to audio-only communication are in the ability of the participants to monitor levels of remote users' attention (Sellen, 1995) and understanding (Olson *et al.*, 1997; Tang and Isaacs, 1993). Participants were also able to detect effects which were confirmed by the process analysis, such as reporting the feeling that turn-taking was difficult. (Olson *et al.*, 1997; Sellen, 1995) therefore providing a form of convergence validity. In particular, users felt able to monitor understanding visually and this seems to mirror results obtained from the coding analysis. One main area where this form of analysis excels, and of potential relevance to the TeachSpeech project, lies in the ability of questionnaires and interview data to monitor user satisfaction levels, as implemented by Anderson *et al.* (1997) in their studies of collaborative problem solving. This information could be especially relevant in highlighting usability issues; problem areas in technology use, changes to the equipment configuration, and future improvements in implementation can all be explored through the opinions of users. Clearly, this is directly relevant to those who take part in the TeachSpeech model of service delivery. Furthermore, similar data could be obtained from the comparison group and in this case, data could gauge opinions and reservations re-

garding potential use of the same technology. This in itself could provide valuable information about how to market such technology to inexperienced users. Finally, as suggested by Olson *et al.* (1997) and Sellen (1995), measuring the perceptions of users allows for one to verify results gained through different forms of analysis such as structural, coding, and performance measures. If consistent, this can add strength to the overall findings of the current study.

1.4.4 Summary

As stated earlier, Monk *et al.* (1996) emphasise the strengths of evaluating mediated communication using a multi-faceted approach. They suggest that the way in which a given form of mediated communication is appraised by a particular evaluation tool will not necessarily reflect the way in which it will be evaluated by other tools; although the potential for convergence validity exists (Olson *et al.*, 1997; Sellen, 1995), it should not be assumed. Thus, this necessitates a host of criteria against which mediated communication should be evaluated in order to provide a full and detailed understanding.

1.5 The TeachSpeech Evaluation

In this chapter, the evaluation methods used for speech and language therapy service delivery models, TM projects, and video-mediated communication were investigated. The aim was to assess the advantages and possible drawbacks of utilising these different evaluation tools in order to design a comprehensive evaluation to apply to the TeachSpeech project. Based on these reviews, the evaluation criteria have been chosen and are grouped into three general categories including the overall performance, the stakeholders' perceptions and the communication process. Each category will be discussed in turn.

Evaluation of the overall performance (hereafter referred to as performance) is comprised of three measures including the clinical effectiveness, a cost analysis and an assessment of how the SLTs spend their time. An evaluation of service delivery models within the speech and language therapy domain illustrated the need for a thorough evaluation of clinical effectiveness. In order to measure the clinical effectiveness, the studies reviewed generally utilised a battery of speech and language tests conducted before and after intervention. Clearly, a measure of clinical perfor-

mance should be utilised in the evaluation of the TeachSpeech project. The most suitable measure in the view of professional SLTs was the Enderby Outcome Measure (EOM). This was selected because it is deemed to be particularly useful as it can assess many different speech and language impairments, which in turn facilitates comparison of impairment severity regardless of the particular type of impairment. While particularly thorough with respect to evaluation of the clinical effectiveness, it was observed that many studies made no other attempts at measuring project performance. In fact, literature in the telemedical field has been criticised concerning its lack of cost analysis studies. In order to counter this criticism, a cost analysis was incorporated into the design of the TeachSpeech evaluation. The cost analysis was derived from a number of measures including log sheets which detailed how the therapist spent their time, and also by gathering information concerning equipment and travel costs. A cost analysis adds crucial information concerning the viability of the TeachSpeech project and is of vital importance to service providers. These log sheets provide yet another measure of performance. An analysis of the log sheets would reveal how therapists spend their time and whether there was a reduction in the amount of travel time in the TeachSpeech model of therapy. These results have implications regarding the productivity and costs of these models of therapy.

The second criterion of interest, namely the stakeholders' perceptions, was demonstrated through a review of the literature which explored the impacts on users of video-mediated communication. Of all the literature reviewed, this was the only one which made thorough assessments of the stakeholders' perceptions; the measures here illustrated the advantage of the video channel. Based on this, questionnaires, and semi-structured interviews were incorporated into the TeachSpeech evaluation to determine the relative merits of the TeachSpeech method of service delivery. These techniques provide valuable information about satisfaction levels with the project, and the confidence of the participants in the ability of the project to deliver speech and language therapy.

The final criterion against which the TeachSpeech model of therapy is evaluated is the nature of the communication process. Previous research on video-mediated communication illustrated that assessing the process of communication can be a useful evaluative tool as it can highlight the advantages or disadvantages of using a video-mediated channel to communicate. It was decided to incorporate

a content coding scheme and structural analysis, to be applied to transcripts of video-mediated communication, as part of the TeachSpeech evaluation to explore the process of communication in this mode of service delivery. An analysis of the dialogues can illustrate features of the interaction of which participants are unaware and can show how the communication link is used during videoconferencing sessions. Unfortunately, because of the lack of comparable face-to-face meetings no conclusions can be drawn as to whether or not any differences in communication are due to the technology or to the increased role of the ESA.

It seems that by examining literature from different fields one gains a unique perspective which enables the design of an extensive and comprehensive evaluation tool. The TeachSpeech evaluation will attempt to encompass the various components outlined above.

1.6 Overview of Thesis

The remainder of this thesis will outline the method by which the TeachSpeech evaluation was implemented and will discuss the results obtained. Chapter 2 will detail the techniques used in the evaluation which appraised the performance of the respective models of therapy and discusses the results obtained. Chapter 3 will outline the results and discussion of the perceptions of the stakeholders. Chapter 4 provides an overview of a small literature concerning different coding techniques, applies a coding scheme and accordingly outlines the results and the discussion concerning the process of the TeachSpeech model of therapy. Finally, Chapter 5 provides the conclusion.

Chapter 2

Evaluating Performance

2.1 Introduction

Conclusions from the various literature reviews in the previous chapter suggest that the performance of the TeachSpeech model of therapy and the traditional model of therapy should be assessed using three parameters. These parameters include clinical effectiveness, an analysis of how the therapists allocate their time and a cost analysis. On the basis of previous literature and the detailed description of the TeachSpeech and traditional model of therapy outlined in Section 1.2.1, three hypotheses were considered. Specifically,

- both models of therapy will be clinically effective in delivering speech and language therapy,
- there will be a significant time saving for the TeachSpeech group, resulting from time saved travelling, and
- there may be a significant difference in the costs of the two models of therapy owing to the small-scale, pilot project and the associated high equipment costs.

2.2 Overview of Chapter

This chapter will examine each of the three parameters used for evaluating performance. The method by which clinical effectiveness is assessed, the results obtained and the discussion that follows are outlined in Section 2.3. Section 2.4 contains the method, results and discussion of how SLTs allocate their time. The cost analysis

in Section 2.5 summarises results presented to the relevant project parties.[1] This analysis includes the costs of both models of therapy in their present form as pilot studies, the projected costs of implementing the models of therapy in their present form as a service, and finally, the costs of implementing the models of therapy in different scenarios as a service. Examples of scenarios investigated include increasing the travelling distance to the schools, and increasing the number of children receiving therapy.

2.3 Clinical Effectiveness

2.3.1 Method

In the following section, we outline the participants in the study, give details of the videoconferencing equipment utilised and provide a summary of the clinical measure used.

Participants

The TeachSpeech group consisted of 12 children who had previously been diagnosed with a speech and language impairment. The children, in both the comparison and experimental groups, were distributed evenly between 4 mainstream primary schools. The TeachSpeech group of children consisted of those children who had received therapy via the TeachSpeech project for 6 months or more and 9 children fell into this category. The remaining 3 children left before receiving 6 months of therapy and therefore their results were discarded. This minimum time period was chosen by the SLT, who suggested that 6 months of therapy would be necessary in order to gauge any reduction of the child's impairment. The following analyses were conducted on the data from the remaining 9 children.

The children's ages at the start of the project ranged from 4 years, 7 months to 5 years, 1 month and their mean age was 4 years, 10 months. Their impairments varied, however, 7 of the 9 children could be described as having specific language impairment. The remaining 2 children were diagnosed with Dyspraxia and Duchenes Muscular Dystrophy. Three of the 9 (33%) children had impairments which contained a significant phonological component.

[1] Final Evaluation Report presented to BT in June 2000

The comparison group consisted of 20 children who had previously been diagnosed with speech and language impairment. Data from 4 children were not available, as one child was only tested once and therefore it was not possible to assess progress, while the other 3 children left before the 6 month cut-off point. That is, children in the comparison group were also required to have undergone a minimum of 6 months speech and language therapy to be included in the analysis. Therefore, the following analysis was performed on data from 16 children. These children's ages at the start of the project ranged from 4 years, 0 months to 8 years, 7 months with their mean age being 5 years, 7 months. As stated previously, the children in the comparison group attended the same mainstream schools as the TeachSpeech children. Seven of the children were diagnosed as having specific language impairment. The remaining 9 children were diagnosed with various behavioural, attentional and specific phonological problems. Again, whether the children's impairment contained a significant phonological component was taken into account and it was found that 6 of the 16 (38%) children fell into this category. Appendix A gives the specific term for each child's impairment in the TeachSpeech group, while Appendix B gives details of the impairments for the comparison group.

Bishop and Edmundson (1987) found that prognosis for children with speech and language impairments differs considerably depending on whether a phonological impairment was present. Those children with a phonological impairment tended to have a better prognosis than children who did not. It has been noted above how many children exhibit phonological problems and since the proportion of children in each group is very similar, it is unlikely that this will be a confounding factor.

The one SLT in the TeachSpeech project was employed by I CAN. When the project commenced there were two SLTs in the comparison group, both of whom were employed by the NHS. One therapist left during the project, leaving only one SLT in the comparison group.

In total 7 ESAs were involved in the project. Throughout the project the number of ESAs varied; 4 ESAs were involved for the entire duration of the project, and thus remained constant, while 3 ESAs used the videoconferencing system to link with the SLT at different points during the project. It should be noted that the ESAs who supported children in the comparison group were the same as those who supported children in the TeachSpeech group.

Equipment

The TeachSpeech model of service delivery involved using a videoconferencing system. As outlined in Section 1.2.1, the SLT and ESA communicated via this videoconferencing link. In Section 1.4, it was demonstrated that different types of videoconferencing systems may affect communication differently, and therefore it is important to detail both how the system was chosen, and which system was utilised in the TeachSpeech study.

A pilot study was carried out to determine the requirements of the technology (Clarke, 1997). Several aspects of the technology were manipulated: the screen size, audio and video bandwidths, lip synchronisation and the video-encoding algorithm. Two I CAN therapists participated in the experiment, in which they were asked to take part in a role-play across the different types of equipment. After each test, they completed a questionnaire. The questionnaire investigated their perceptions about the image quality, audio quality and how the technology affected the communication process. In total, thirty-one test sessions were run. Once all the test sessions were complete, a final questionnaire was administered. This related to whether or not the participants had taken part in a test session which in their opinion had sufficient audio and video quality for their purposes. The results showed that the BT VC6000 videoconferencing system, using a digital bandwidth of 384 kb/sec formed by 6 basic rate channels at 64 kb/sec (i.e. 3 ISDN2 lines) was a suitable option. This is a high quality, mobile 'rollabout' system, comprising a 48 kilobit audio channel, a 28 inch colour monitor and camera capable of pan, tilt and zoom, remote comparison and picture in picture facility. As ISDN video only transmits the parts of an image that it thinks have changed since the last frame, a picture remaining mainly motionless, for example, a head and shoulders perspective of an SLT talking, the frame rate is much higher than when transmitting a moving picture. The frame rate of the video in the TeachSpeech project can be categorised in the 6.25 to 12.5 frames per second band.

One videoconferencing unit was located in the I CAN Nursery and the remaining four were situated in schools. Two of these schools were in the immediate local area and the other two schools were approximately 15 and 7 miles away. The videoconferencing units in the mainstream schools were all situated in staff rooms.

The I CAN nursery and three of the four schools had very few technical

problems, but the fourth school experienced considerable technical difficulties. The BT engineers believed that the problem was a failure in the termination of the cables in the wall sockets that the VC6000 used, resulting in an unreliable electrical connection. This type of problem is very unusual with this equipment and it is also particularly difficult to resolve because it is intermittent. It affected how frequently the SLT and ESA linked up, although when a link was made, the quality of the videoconference was comparable with the other schools which were taking part in the project and therefore the data from this school were still included in the analysis.

Clinical Measure

EOMs were used to assess each participant. The EOM is a diagnostic instrument which was chosen by the I CAN therapist and was deemed to be particularly useful as it can assess many different areas of a particular speech and language impairment. It also allowed comparison between different speech and language impairments. Children were rated on five different factors: impairment, disability, handicap, patient well-being and caregiver well-being. A scale of 0 to 5 was used to rate each factor. For example, when rating 'impairment', a value of 0 would represent 'the most severe presentation of this impairment' and a value of 5 would represent 'no impairment'. For 'disability' a value of 0 represented 'totally dependent unable to function' and 5 represented 'independent able to function'. When rating 'handicap', 0 represented 'no autonomy, isolated, no social / family role' and 5 represented 'integrated, valued, occupies appropriate role'. A 0 for 'well-being' for either the caregiver or patient represented 'severe constant upset / frustration / anger / distress / embarrassment / concern / withdrawal' and 5 represented 'no inappropriate upset / frustration / anger / distress / embarrassment / concern / withdrawal'. A brief paragraph was provided for the therapist which outlined the appropriate criteria for each rating scale in each factor.

Reliability of Clinical Measure

As two therapists conducted the assessments using the EOMs, a test of inter-judge reliability was required, to test whether the therapists rate the children in a similar manner.

Towards the end of the project, in July 1999, the two therapists picked ten

Table 2.1: Recommended benchmarks for the interpretation of Kappa values

Kappa Statistic	Strength of Agreement
< 0.00	Poor
0.00 - 0.20	Slight
0.21 - 0.40	Fair
0.41 - 0.60	Moderate
0.61 - 0.80	Substantial
0.81 - 1.00	Almost Perfect

recent files from the NHS database. Both therapists established that the patients were not known to them. The process of rating each file involved examining each file at length, which took approximately 15 to 25 minutes. The therapist would then rate the patient at the initial and final stages of treatment using the EOMs. This process was completed for ten files. Since each EOM produces 5 data points and the measure is conducted twice on each file, 10 data points are produced for each file. From 10 files 100 data points are produced, and out of these 100 judgements, 47 were found to be in agreement with an agreement of 47%. The Kappa coefficient of agreement (Siegel and Castellan, 1988) was calculated and a reliability coefficient of 0.32 was obtained. Landis and Koch (1977) recommend an interpretation of the Kappa coefficients; see Table 2.1. A Kappa of 0.32 indicates 'fair' agreement, although this is at the low end of the scale. To investigate the results further, the data were divided into the different components that the EOM investigates. Kappas were calculated for each component. Impairment 0.15, Disability 0.30, Handicap 0.16, Patient Well-being 0.39, Caregiver Well-being 0.54.

If the therapists rated the children differently, this would have serious repercussions. Any conclusions made concerning the clinical effectiveness would be highly unreliable. Therefore further investigation was carried out in order to determine whether or not this was the case. The possibility exists that either a type I error or a type II error might have occurred; a type I error is where one accepts the hypothesis that the therapists rate the children differently when in fact they rate them similarly, while a type II error involves rejecting the hypothesis that the therapists rate the children differently when in fact they do. In order to limit this possibility, two adjustments were made. Firstly, it was decided that as the amount of information in the files was extensive, this should be limited to key reports in order to allow

therapists to focus on specific information provided. Secondly, we ensured that each therapists had the definitions of each EOM placed in front of them. This was done so that they could refer to the definitions whenever they wished. The new procedure was repeated using 10 different files. Out of a possible 100 judgements, 72 were found to be in agreement, giving a percentage agreement of 72%. The Kappa coefficient of agreement was calculated and a reliability coefficient of 0.63 was obtained. The significance of the Kappa was calculated and showed that the agreement was very unlikely to be due to chance factors ($p<0.01$). Substantial agreement was found, and this is a critical finding. If the therapists had assessed the children differently, the findings concerning the clinical effectiveness would have been considered tenuous at best. Since the findings indicated substantial agreement between the therapists, the following results can be considered credible.

2.3.2 Results

Enderby Outcomes were used to assess the clinical severity of the children's impairments. The raw data from the Enderby Outcomes were considered as ordinal data and therefore non-parametric statistics were used. A Mann-Whitney Test for independent measures design was performed to investigate whether or not the two groups were significantly different at the first and final assessment. It was found that the first assessment data did not significantly differ. The same statistic was conducted for the final assessment and again the two groups did not differ significantly.

Table 2.2 illustrates the medians of the two models of therapy at each assessment time. Each group significantly improved over time on each factor and the groups were statistically equivalent at the first and final assessment.

A Wilcoxon Signed Ranks Test for repeated measures design was used to see if the TeachSpeech group significantly differed on any of the factors between the first and last assessment. Results suggested each factor had significantly improved. Results are illustrated in Table 2.3.

The same statistic was performed on the comparison data and the results also showed a significant improvement for this group as shown in Table 2.4.

2.3.3 Discussion

To reiterate, the first hypothesis was

Table 2.2: Median scores of first and final assessments on the Enderby Outcome Measures for both conditions

First Assessment	TeachSpeech Group	Comparison Group
Disability	3.00	3.00
Handicap	3.00	3.00
Impairment	3.00	2.75
Well-Being of Caregiver	3.00	3.00
Well-Being of Patient	3.00	3.00
Final Assessment	**TeachSpeech Group**	**Comparison Group**
Disability	4.00	4.00
Handicap	4.00	4.00
Impairment	4.00	4.00
Well-Being of Caregiver	3.50	4.00
Well-Being of Patient	4.00	4.25

Table 2.3: Results of EOMs at first and final assessment for the TeachSpeech group. An asterisk indicates a significant difference to $p<0.05$.

EOM component	Z score	Significant Difference to $p<0.05$
Disability	Z=3.09	*
Handicap	Z=2.96	*
Impairment	Z=3.07	*
Well-Being of Caregiver	Z=2.83	*
Well-Being of Patient	Z=2.84	*

Table 2.4: Results of EOMs for first and last assessment of the Comparison Group. An asterisk indicates a significant difference to $p<0.05$

EOM component	Z score	Significant Difference to $p<0.05$
Disability	Z=2.54	*
Handicap	Z=2.59	*
Impairment	Z=2.53	*
Well-Being of Caregiver	Z=2.20	*
Well-Being of Patient	Z=2.59	*

- both models of therapy will be clinically effective in delivering speech and language therapy.

From the clinical data gathered, it can be seen that for the initial and final assessment both groups were equal, both in terms of the severity of their impairments and their improvement over time. The EOM which assessed the severity of a child's impairment, including the level of the child's and caregiver's well-being, suggested there are no differences in the clinical advancement of the children between the two programmes, and thus both models were equally effective.

Several points needed to be raised concerning the limitations of the evaluation of the performance of the models of therapy. Firstly, the limited subject population was clearly a drawback. Ideally a larger population with similar impairments could have been utilised in the study. However, considering the setting of the study, there was little flexibility in the number of children available. Bishop and Edmundson (1987) have found that children with phonological disorders have a better prognosis than most other speech and language impairments. The proportion of children, in this study, whose impairment contained a phonological component, was noted in Section 2.3.1. Since the findings showed that both groups had relatively equal numbers of children who exhibited a phonological impairment, this should not confound the results.

Secondly, it would have been ideal to measure clinical efficiency using a variety of measurements. Ruscello *et al.* (1993) assessed how two different methods of service delivery affected children with phonological disorders. The measurements taken to assess the treatments consisted of the Khan-Lewis Phonological Analysis (1986) test, which consists of a 30 item task and a questionnaire. Ruscello's conclusions regarding the clinical efficiency of each model of service delivery was derived from the amalgamation of the results obtained from all of their measures. Due to practical reasons, which were mainly economic notably the cost of therapists' time, the EOM was the only method of assessment used to evaluate the progress of the children. This measure is particularly advantageous since it assesses a wide variety of factors including the well-being of the patient and the caregiver. If possible, it might have been useful to administer a battery of tests to ensure coverage of a wider range of factors.

Thirdly, it is important to note that the statement 'both groups improved significantly over time' is not interchangeable with 'clinically significant change'. To stipulate that there was 'clinically significant change' or that the therapy *per se* was effective, the study would need to demonstrate that the change was not due to maturation, that the measurements used were reliable and valid, that the change was of a reasonable size (at least one standard deviation of the measure used), and the change impacts on the patient (Bain and Dollagham, 1991). The study conducted here meets two out of the four criteria. Firstly, the measurement used was reliable and valid (Enderby, 1997) and secondly, impact on the patient of the clinical change that takes place was considered. Bain and Dollagham (1991) suggest establishing a multiple base line for various behaviours to allow the monitoring of language changes resulting from treatment and maturation. Since this study is concerned with the comparison of two models of therapy and does not contain a non-intervention group, no claims are made regarding clinically significant change or the effectiveness of the therapy *per se*.

Could the clinical outcome have been influenced by the time each child spent with an ESA? Unfortunately we are not able to provide a definitive answer as the amount of time the ESA spent with the child was not logged. At the onset of the project, this measure was not considered necessary since the SENCO decides how much time an ESA spends with a particular child in a standardised manner, which is not wholly related to the child's speech and language impairment and may depend on other factors such as behavioural problems or learning difficulties. Considering the children in the different groups attended the same schools, and were therefore under the direction of the same SENCOs, it was felt that this measure was not necessary. In retrospect this measure would have been useful and could have consequences regarding the clinical results. It would be recommended that any follow-up study include this measure.

2.4 Allocation of Therapists' Time

One of the main aims of telemedical projects is to reduce the amount of time the specialist has to travel in order to increase clinical productivity and patient turnover

rates. It is important to detail any differences regarding how the therapists spend their time in the TeachSpeech project as compared to the traditional model of therapy.

2.4.1 Method

In this section, we aim to outline the participants involved and the method used in assessing how the therapists allocated their time.

Participants

At the start of the project there were 3 SLTs, however only 2 SLTs took part in the entire project. Thus, the primary SLTs, included the only TeachSpeech SLT, and the single NHS SLT who was part of the study throughout the project. These 2 SLTs served as the participants.

Log Sheets

The aim of the log sheets was to record which task the therapist completed with the child and the amount of time this task required. From the log sheets estimated costs could also be calculated.

A log sheet was kept for each child. The categories on the log sheets included face-to-face contact, liaison time with others, preparation time, training time, travel time and administration time. To better understand these categories, each will be described in turn. Face-to-face contact was noted when the SLT spent face-to-face time with the child, either in the form of providing speech and language therapy to the child or when the child was being clinically assessed. Liaison time with others was recorded when the SLT consulted others regarding a child. For example, telephone calls to parents or teachers and the videoconferencing sessions between the SLT and ESA would be included in this category. Preparation time was any time the therapists spent preparing for a session with a child, teacher, parent or ESA. Training time included time when the SLT received advice or training on a particular speech and language impairment. This might include communicating with another professional such as a doctor, or it might be a specific training day that the SLT attended. Travel time was defined as any time the therapist spent travelling from their place of work to visit a child, parent, teacher or ESA. Administration

time was recorded when the therapist spent time dealing with organisational issues. Examples include writing reports, making telephone calls to arrange meetings, or writing letters.

The two principal SLTs filled out the log sheets as time progressed and they were collected at the end of each term. It was noted through the interviews that the log sheets were filled out approximately once a week and this was true for both SLTs. It is difficult to estimate the accuracy of the log sheets because they were being completed weekly and the therapists were estimating the amount of time spent on each task. Thus, this type of measurement is subject to a certain degree of human error and the log sheets should therefore be considered with this in mind.

2.4.2 Results

Independent t-tests were conducted to investigate whether there were any significant differences between the TeachSpeech model of therapy and the traditional model of therapy. Prior to the actual analysis, it was expected that the TeachSpeech SLT would not spend as long on each child and that this time saving would be a result of time saved travelling.

There were no significant differences between the two conditions regarding the total amount of time the therapist spends on the child. The total amount of time was divided into the different categories as noted in the log sheets. Independent t-tests were calculated for each of the sub-components. It was found that, in the traditional model of therapy, the therapist spent significantly more time travelling $t(17.327)=2.225$, $p<0.05$, in face-to-face contact with the child $t(17.239)=7.223$, $p<0.05$ and significantly more time preparing $t(16.678)=3.793$, $p<0.05$. Significantly more liaison time $t(8.912)=6.171$, $p<0.05$ occurred in the TeachSpeech group. The mean average time per month set aside for each child and the breakdown of how this time is spent has been calculated and is shown in Table 2.5.

Table 2.5 shows that, although the two groups did not differ in the total amount of time allocated to each child overall, there were differences between the sub-components measured. As can be seen from the table, the SLT in the Teach-Speech group did spend face-to-face time with the child, mainly for assessment purposes.

Table 2.5: The mean time in minutes and the corresponding standard deviation (SD) set aside each month for each child, by the SLT. An asterisk in the results column indicates a significant difference between the two groups, $p<0.05$.

Time Allocation	TeachSpeech		Comparison		Results
	Mean	SD	Mean	SD	
Face-to-face contact	6.2	1.9	18.6	6.6	*
Liaison with others	32.3	11.8	7.5	3.7	*
Preparation time	0.3	0.5	2.9	2.5	*
Training time	0.2	0.4	0.5	0.7	
Travel time	6.0	4.3	16.7	18.5	*
Administration time	12.3	10.4	14.8	7.5	
Total	57.3	18.9	61.0	31.8	

2.4.3 Discussion

The results showed that there was no difference between the two models of therapy in the average amount of time per month a therapist spent on a child. However, when the sub-components of the log sheets were investigated, differences were found.

The TeachSpeech group did not spend as much time travelling as the comparison group, saving on average 10.7 minutes per child per month. As stated previously, of the four schools using the videoconferencing systems two were situated locally while the remaining two were located approximately 15 and 7 miles away. It is worth noting that if all the schools chosen had been located at more remote distances, the difference in travel times between the two models of therapy might have been such as to have affected the overall time spent on the children for each model of therapy.

Time savings also occurred in the TeachSpeech group for preparation time (2.6 minutes per child per month) and face-to-face contact with the child (12.4 minutes per child per month). In the TeachSpeech group, the therapist mainly spent time liasing with ESAs across the videoconferencing link. The SLT in the comparison group spent more time in face-to-face contact with the child, in travelling and in preparation. Two sub-components which did not significantly differ were the time spent completing administration duties and training time.

It might be suggested that the therapists allocated their time evenly to each child, irrespective of the child's needs. If this was the case, little could be said regarding the use of the therapists time under the different models of therapy. There

are, however, two pieces of evidence which counter this suggestion. Firstly, the standard deviations presented in Table 2.5 showed that the way therapists spent time varied greatly within categories and this relatively high standard deviation is also present for the total amount of time. This illustrates that although the means are similar, the amount of time spent on each particular child varies considerably. Secondly, therapists, when canvassed using questionnaires and interviews, stated that they allocated their time according to the child's needs. Thus, an analysis of how the therapists distributed their time is particularly informative for the TeachSpeech evaluation.

It has been established that the two models of therapy do not differ in the amount of time allocated for each child, but there were differences in how the time was spent. One of the main goals of telemedicine is to reduce travel time, which would release the expert to use their time more effectively. The TeachSpeech project certainly did this, even though the schools are not located far from the base of the SLT. The TeachSpeech model of therapy also reduced the amount of time the therapist needed to spend in preparation. Evidence from interview data suggested that the saving in preparation time was due to 'the feeling that everything is at hand'. If the therapist needed any tools during a session with an ESA she could retrieve them with ease.

It is difficult to quantify which uses of the therapists' time are more valuable. However, it could be suggested that face-to-face contact with the child and liaison time are more valuable than travel or preparation time. This is especially the case with travel time because, for example, therapy-related activities utilise the specialist skills of the therapist, while travelling time does not. The difference, if there is one, between the value of face-to-face contact and liaison time is more complex. Face-to-face time, where the therapist assessed or treated the child in a one-on-one scenario might at first seem more valuable, but other issues need to be taken into consideration. The liaison time that occurred in the TeachSpeech project continued over a sustained period of time with the ESA and this had several benefits. Firstly, the therapist could modify the treatment programme to the specific needs of the child on a weekly basis. Secondly, if needed, the therapist can provide support to the ESA regarding treating unusual impairments. Finally, the ESA was likely to become more confident and competent supporting the child, and support for this

comes from the interview and questionnaire data.

The perceptions of the stakeholders have also provided insight into this area. The interviews with the TeachSpeech SLT highlighted her ability to change treatment programmes when necessary, being able to deal with new speech and language problems a child was having almost instantly and being able to provide support to the ESA regarding unusual problems. All the interviews conducted with the ESAs asked them to compare their confidence levels in supporting the TeachSpeech children as opposed to the comparison children. Of the eight interviews conducted with ESAs, six stated that their confidence was higher when dealing with the TeachSpeech children while the remaining two stated that confidence levels were comparable. This finding seemed to stem from the fact that expert advice relating to children in the TeachSpeech group was being delivered in a formal and frequent manner to the ESAs, which helped them in their work with the children. The value of having ESAs who were more confident in supporting children with speech and language impairments could be considerable. Generally ESAs spend time with the child on a daily basis, and thus contribute a great deal to their treatment programme.

Overall, measuring how the therapists spend their time has proved a useful measure. Differences in the way in which therapists apportioned their time was found and these differences may have repercussions, for example, in terms of cost.

2.5 Cost Analysis

Service providers are clearly interested in the respective costs of each model of service delivery. A cost analysis is therefore of practical importance and contributes information regarding the measure of performance. In the following we summarise results as presented to BT as part of the final evaluation report of the TeachSpeech project (Katsavras, 2000).

2.5.1 Method

Cost analysis research can be conducted from a variety of perspectives (Eiserman *et al.*, 1990). Consequently, the first step in a cost analysis is to state which perspective the analysis is taking and to include or exclude the appropriate costs (Field, 1996). Here the analysis has been calculated from the viewpoint of the organisation delivering the therapy. The second step is to determine the costs involved in each model

by identifying all the resources necessary to operate each of the models of therapy (Levin, 1983). The resources costed for the therapies here include: fixed costs, such as personnel and equipment; direct variable costs, such as travel and hourly telecommunications cost; and indirect variable costs, including insurance, postage and marketing (Allen and Stein, 2000). Personnel and travel costs were determined from the log sheets. This was done by transforming the total average amount of time per child into salary costs and by translating travel time into miles travelled. It is also worth noting that this analysis was not a cost-effectiveness analysis, which is defined as an analysis which directly compares the cost of a model of therapy to the clinical outcome (Cukier, 1997). This analysis was a simple cost investigation that, taken together with the clinical effectiveness results and the amount of time spent by therapists on each child, could be used to make decisions regarding relative merits of the two different models of therapy. The method of calculating the costs was derived following advice from a range of professional bodies such as the Royal College of Speech and Language Therapists and NHS Salisbury District Hospital. Full details are provided in Appendix C

2.5.2 Results

The Cost of Implementing the Pilot TeachSpeech Project

The most common method for evaluating the cost of studies is by comparing an approach with at least one other (Cukier, 1997). This type of investigation has been carried out here. Two cost analysis outlines have been provided. The first includes the costs to run either the TeachSpeech or comparison model of therapy for 3 years, on 9 children, at 4 different schools, using 5 videoconferencing terminals. The relevant details are shown in Table 2.6. Allen and Stein (2000) defined the three categories involved in cost analysis, these are fixed, direct variable and indirect variable costs. The costs from both models of therapy have been classified accordingly. For a detailed breakdown of the costs see Appendix C.

It should be noted that the above costs were for implementation of the TeachSpeech trial only and that this implementation can be satisfied in several other ways that may be cheaper. In the future these costs could increase or decrease depending on how prospective implementations are built and run. For example, the Sony Contact 29/384 which is claimed by the manufacturers to be broadly similar to the

Table 2.6: Resources and the costs for each model of therapy. For more details see Appendix C.

Resources	TeachSpeech	Comparison
Fixed costs		
SLT - £21,743 per annum	£3,234	£3,443
Videoconferencing units - £7,950 per machine	£39,750	N / A
Start up costs - £199 per machine	£995	N / A
Direct variable costs		
Travel - 39.8 pence per mile	£388	£887
Line charges - £380 per quarter	£22,800	N / A
Indirect variable costs		
Insurance - £200 per machine per year	£3,000	N / A
Total cost	£70,153	£4,330

BT VC6000 is £1,000 cheaper per system. There are other systems in use which are of lower specification and they may be acceptable for telemedical projects such as the TeachSpeech project but their quality of fit to the requirements is untested. Examples of such systems include the BT VS1 and the NetView LAN from Global Communication Solutions. These PC-based systems use only 2 basic rate channels (i.e. one ISDN2 line) and are cheaper to buy at approximately £5,000 per system.

The Cost of Implementing the TeachSpeech and Traditional Therapies in their Present Forms as a Service

Perhaps a more useful comparison might be to compare the costs of implementing both models of therapy as a 'service'. For this, certain key considerations need to be made; the number of children who could benefit from one terminal, the distance and therefore length of time it would take to travel to the schools and how long the service would last. Under the assumption that there are four schools, to calculate how many children might benefit from a video link from one school, we need to examine how many children from the general population require speech and language therapy. One in 20 children in schools required speech and language therapy,[2] and there were on average 166 pupils in a primary school in the Wiltshire area.[3] Based on these figures, a school might well have 8 children requiring speech and language therapy.

The second consideration concerns the schools. Schools that would benefit

[2]Royal College of Speech and Language Therapist, 1998
[3]Wiltshire Education Authority Booklet, 1999

Table 2.7: Cost of Delivering the TeachSpeech and Traditional Model as a Service. For more details on how these costs were derived see Appendix D

Resources	TeachSpeech	Comparison
Fixed costs		
SLT - £21,743 per annum	£24,653	£31,945
Videoconferencing units - £7,950 per machine	£39,750	N / A
Start up costs - £199 per machine	£995	N / A
Direct variable costs		
Travel - 39.8 pence per mile	£12,838	£29,344
Line charges - £380 per quarter	£38,000	N / A
Indirect variable costs		
Insurance - £200 per machine per year	£5,000	N / A
Total cost	£120,241	£61,289

most from the remote link would be situated at least a moderate distance away from the SLT and therefore would gain from benefits associated with reduced travel time and travelling costs. Here it was assumed that each of the four schools was situated 30 miles from the SLT and that a 60 mile return journey would take 80 minutes to complete. The average number of times per month the SLT, in both conditions, visited the schools was determined from the log sheets. To determine this, children who travelled to the clinic for treatment were excluded. For the cost analysis of the service delivery model, these figures were translated into the mean number of miles per month a therapist would travel. From this, the travel costs could be calculated. The salary costs were also adjusted.

The remaining consideration concerned the length of time the service would run. It was assumed that the service would continue for longer than three years and in this instance the costs will be calculated over a five year period as a longer time period may well involve replacing equipment. While the equipment costs which have been used here were the same as those in Table 2.6, trends have demonstrated however that costs are likely to fall (Cukor *et al.*, 1995).

Table 2.7 shows the resulting costs if the TeachSpeech trial implementation were scaled up for service delivery. The results show that in moving from a trial to a service implementation, it would be desirable to use this transition as an opportunity to drive down implementation costs because, although the TeachSpeech project is still more expensive, the gap between the two services is narrowing and thus reducing the cost of the technology may result in a cost saving for the TeachSpeech group.

The Costs of Various Implementation Scenarios of the TeachSpeech Concept

A third cost analysis has been conducted below. This altered various components, such as the distance of the schools, the number of children being supported and the lifetime of the equipment. The graphs below illustrate where the TeachSpeech project becomes economically efficient. This cost analysis uses, where appropriate, the same assumptions as outlined in the previous section. It should be noted that the cost analysis here used the BT VS1 videoconferencing system. As discussed above, the cheaper VS1 technology may be more appropriate than the VC6000, although whether or not the VS1 fits the requirements of the project would have to be assessed.

Travel Time

Figure 2.1 illustrates how the cost of both models of therapy increased as the location of each of the four schools becomes more remote. The cost analysis has been calculated assuming that 32 children would utilise the system and that the VS1 unit would have a life span of 5 years. It can be seen that, when the schools are located over 50 miles from the SLT's location, the TeachSpeech model became cost effective. From the gradient of the line it is demonstrated that for every additional mile the SLT has to travel, it cost the TeachSpeech group £326 over the 5 year period for the 32 children. In the comparison group, the same calculation results in a cost of £746 per additional mile.

Number of Children

It is interesting to investigate how the numbers of children utilising the system affect the cost of the respective models of therapy. It may well be the case that other health and educational services can tap into the TeachSpeech service and thus the numbers of children using the system would increase. Figure 2.2 shows the results of the cost analysis when larger groups of children were being supported by the system. The gradients were examined. They demonstrated over a 5 year period, that the cost per additional child who was located in a school 40 miles away was £2328 for the comparison group and £1352 for the TeachSpeech Group. It is clearly cheaper to utilise the TeachSpeech set-up if more than 45 children are being supported.

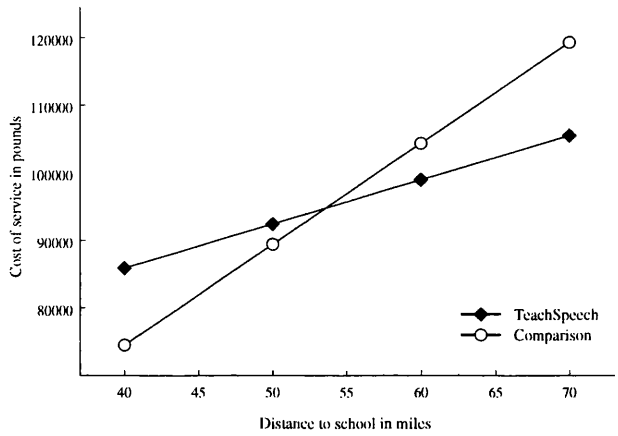


Figure 2.1: This figure represents the increasing total costs of the respective models of therapy, as the distance between the SLT and the location of the schools increases.

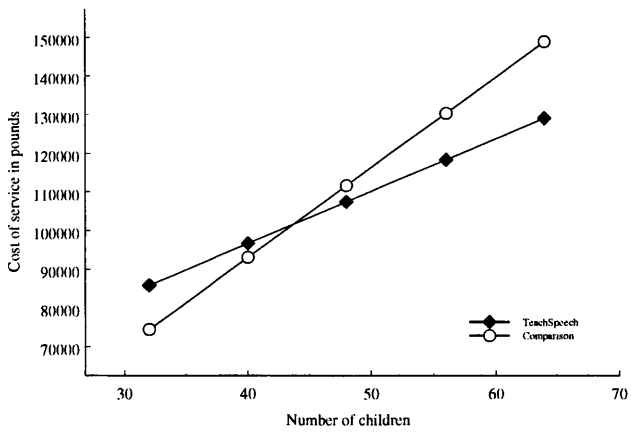


Figure 2.2: Increasing the number of children being supported by the TeachSpeech and traditional models of therapy.

Life Span of Equipment

Figure 2.3 illustrates how the life span of the equipment affects the overall cost analysis of the two models of therapy. The cost analysis was calculated in accordance with the previous assumptions, where it is assumed that there are 32 children receiving support located in schools situated 40 miles from the SLT. It can be seen that over a period of approximately 9 years, the TeachSpeech group costs less than the comparison group. The gradients illustrate that the TeachSpeech group cost £11,687 per year for each additional year, and the comparison group cost £14,899 per year for each additional year. It takes 9 years to recuperate the initial outlay of the videoconferencing equipment. The time scale to deliver this reduced cost may be undercut by changes in the technology rendering them obsolete or by organisational policy which requires the equipment to be replaced.

Combination of all Three Factors

The final step was to investigate how the combination of all three factors would affect the crossover point of the two lines. Figure 2.4 represents a combination of Figures 2.1, 2.2 and 2.3. It can be seen from the graph that, as the distance of the schools increased, the number of children utilising the system increased and the life time of the equipment increased, the TeachSpeech project became increasingly cost effective. This graph demonstrates which scenarios were cost effective for the respective models of therapy. However, it is unlikely that all factors can be taken into consideration simultaneously. For instance, in such areas as the Highlands and Islands, although the distances are large the number of children requiring support is likely to be small due to low population densities.

2.5.3 Discussion

The cost analysis has demonstrated that the TeachSpeech pilot study costs more than the comparison group. A cost analysis for the delivery of a service was also calculated. Here too, the cost of the TeachSpeech model was higher. Finally, a cost analysis was conducted in various scenarios. It was found that in certain circumstances the TeachSpeech study was less expensive than the traditional model of therapy. There are several issues to be taken into consideration, and each will be discussed in turn.

The first issue concerns the cost of buying and maintaining the equipment.

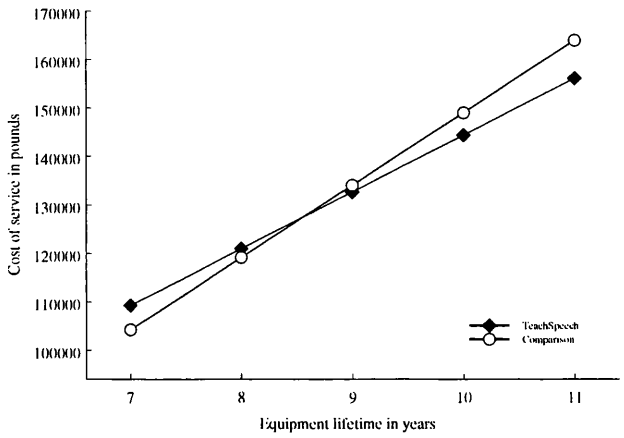


Figure 2.3: Represents how increasing the life span of the equipment affects the total costs for both models of therapy.

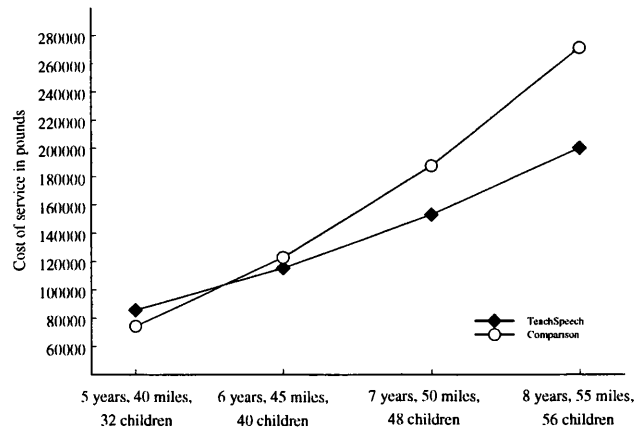


Figure 2.4: Represents how the combination of all factors presented in the previous figures affects the total costs of both models of therapy.

The equipment used for the TeachSpeech project was taken from the higher end of the market in terms of cost and quality. A pilot study was conducted to determine which equipment would be most suitable (Clarke, 1997) and considered the possibility that the children would speak directly with the therapist over the link. As this functionality was not required in the pilot study, it may well be feasible to use systems which are cheaper but still support the type of interaction which took place across the link, such as the VS1 Compact Videoconferencing System (£5,000) or the Sony Contact 29/384 (£7,000). The VS1 is a more basic system with lower video and audio quality. The Sony Contact is very similar to the VC 6000 and costs nearly a thousand pounds less. Maintaining the equipment is another source of cost for the TeachSpeech project. When calculating the cost for the service delivery model, the benefits of reduced travel time are not being realised due to the high operating costs. Often the lower quality systems only use 3 ISDN lines and are therefore cheaper to maintain and operate. It is worth noting that these operating costs are constantly decreasing (Cukor *et al.*, 1995). Therefore, in the future the TeachSpeech model of therapy might become more feasible.

Secondly, the possibility of using the system for other functions exists, and this has consequences for any cost analyses conducted. As part of the demonstration programme which canvassed prospective users, professionals from thirteen different areas were asked their views as to whether the technology would be useful for their specific profession. Results showed that the vast majority of the professionals viewed the system as useful. This suggests that other professions, such as clinical or educational psychology, could share costs of using the videoconferencing system which would call for a re-calculation of the cost analysis.

Finally, as raised earlier, in the TeachSpeech model of therapy the SLT spent time with the ESA, providing support and tuition. The ESAs generally have had no formal training in speech and language therapy techniques and thus this support and tuition is invaluable. The long term benefits of training the ESAs may be considerable as the ESAs can use their knowledge to support many children in their present and future care and the ESAs felt more confident in the support they gave. The monetary value of this has not been taken into account since it is difficult to quantify.

In summary, various considerations need to be taken into account in order

to calculate the costs of the models of therapy and it was shown that in specific circumstances the TeachSpeech model can be cheaper than the traditional model of therapy.

2.6 Conclusions

From the evaluation data, and through examining the performance of the models of therapy, several conclusions can be drawn. The EOMs highlighted that there were no differences between the clinical progress of the two groups of children. This was an important finding, since this showed that the innovative TeachSpeech model of therapy was as clinically effective as the traditional model of therapy. The clinical effectiveness findings are in line with much of the previous research outlined in Section 1.2.2.

The TeachSpeech SLT did save travel time and although this does not result in an overall time saving, it potentially allows the therapist to use their time on more productive activities. An extensive cost analysis, which encompassed a review of how the therapists spent their time, demonstrated that the TeachSpeech model was more expensive, but this difference may well diminish as operating costs are liable to fall with the development of more affordable technology in the future.

Overall, the results from the three performance measures of the respective models of therapy were complex and it was not obvious which model ought to be favoured. The results, however, clearly showed that the use of more than one measure to assess the performance was a prudent step and that the multi-faceted evaluation of performance has proved worthwhile.

Chapter 3

Evaluating Perceptions

3.1 Introduction

The second evaluation criteria used in this study concerns the perceptions of those involved in both the TeachSpeech group and the comparison group. Extensive evaluation was undertaken which compared the two models of therapy in terms of performance. In Section 1.2.2, Section 1.3.2 and Section 1.4, it was shown that previous studies have also included the evaluation of stakeholders' perceptions. In this study, questionnaires and semi-structured interviews were used to assess the perceptions of the stakeholders. A hypothesis was developed which concerned the perceptions of the stakeholders. Specifically,

- stakeholders will be equally satisfied with both models of therapy.

3.2 Overview of Chapter

This chapter examines the perceptions of the stakeholders. Section 3.3 reviews the tools used to capture the attitudes of the stakeholders. Section 3.4 outlines the results obtained and Section 3.5 provides a discussion. Finally, Section 3.6 presents the conclusions. Included in this chapter are the attitudes of prospective users who are not directly related to the project, yet their opinions, as captured by questionnaires, are useful in assessing the potential utilisation and likely success of this method of therapy delivery. The method, results and discussion regarding this are integrated into the appropriate sections.

3.3 Method

3.3.1 Stakeholders

In the TeachSpeech group, those who were considered stakeholders were canvassed for their opinions, including the SLT and parents of the speech and language impaired children. The comparison group stakeholders consisted of the two SLTs and the parents of the children. Stakeholders who spanned both groups included ESAs, teachers and head teachers. Professionals in the health and education sector who observed a demonstration of the technology were also canvassed using a questionnaire. To be consistent, the term stakeholder has been utilised throughout this chapter to describe the above-mentioned parties.

3.3.2 Questionnaires

Questionnaires were sent to all the stakeholders. For those who utilised the video link, the questionnaires assessed the stakeholders' views of the usability of the technology, satisfaction levels with the process, outcomes of the videoconferences and their prior computer experience. The questionnaires used in the evaluation were based on the Multimedia Communication Questionnaire which focused on usability of videoconferencing systems. Various sections of the questionnaire used in the evaluation were specifically designed to investigate user satisfaction and the working environments of the SLTs and the ESAs. Questionnaires given to the parents mainly assessed satisfaction levels. The majority of the questions were constructed using a five-point Likert scale where the midpoint represents a neutral response. The questionnaires were anonymous, but required respondents to give occupation and date of birth. The questionnaire sent to the comparison group explored many of the same topics, with the exception of questions concerning the quality of the videoconferencing technology. It did however include a section concerning the prospect of utilising videoconferencing as part of the model of speech and language therapy.

The first sets of questionnaires were administered in December 1997. These were issued to the SLT, the parents and the ESAs in the TeachSpeech trial. The second sets of questionnaires were issued to the comparison SLTs and parents in January 1998. Follow-up questionnaires were administered to the TeachSpeech group approximately a year later.

The I CAN Nursery also provided a demonstration programme whereby relevant parties were invited to learn about the TeachSpeech project and to find out how videoconferencing technology worked. As part of the demonstration, those invited were asked to fill out a questionnaire which explored how they felt about the technology and whether they felt such technology would be valuable for their own profession. These stakeholders were asked to complete the first part of the questionnaire before the demonstration and the second part after the demonstration so that attitude changes could be determined. Questionnaires given to those who were directly involved with the technology addressed issues such as satisfaction levels and user needs. The demonstration questionnaires examined the wider picture, including an investigation of how others perceive the technology and the system in which it is embedded.

3.3.3 Interviews

Semi-structured interviews were carried out in March 1998, November 1998 and May 1999. All stakeholders were invited to attend an interview and those interviewed included stakeholders from each group. The focus of the interviews differed where tailored questions were compiled from relevant literature and experience. In relation to the SLTs, ESAs and teachers, the main aim was to examine the stakeholders' work experience, their working environment and whether it had been affected by the TeachSpeech model of therapy, what direct comparisons between the two models of therapy could be made, and user satisfaction levels. The parents were asked if they felt the system was effective, their opinions on the effect of the technology on the well-being of their children, and any negative or positive feedback they had about the system. The interviews were carried out in the appropriate school or home of the stakeholder and were all audio recorded. In total, 29 stakeholders were interviewed.

3.4 Results

3.4.1 TeachSpeech Group - Questionnaires

The first sets of questionnaires were sent to the TeachSpeech SLT, ESAs and parents in December 1997. At the beginning of January 1999, the second or follow-up questionnaire for the SLT and ESAs were sent out. The parents also received a second

Table 3.1: TeachSpeech questionnaire completion rates where the numerator represents data corresponding to the first questionnaire and the denominator corresponds to the second.

Category	Category Population	Number of Respondents	Percentage of Respondents
SLT ($1^{st}/2^{nd}$)	1/1	1/1	100/100
ESA ($1^{st}/2^{nd}$)	7/7	6/4	86/57
Parents ($1^{st}/2^{nd}$)	9/9	8/5	89/56

questionnaire towards the end of February 1999. The completion of the second set of questionnaires was timed to be approximately one year after the completion of the first set of questionnaires. Response rates are indicated in Table 3.1. The material from the questionnaires is presented below. Data that are deemed to be of particular interest are shown in this section. The full results of these questionnaires are given in Appendix E (SLT), Appendix F (ESAs) and Appendix G (parents). The results from the first and follow-up questionnaire have been amalgamated; where appropriate, the results have been reported in isolation.

Speech and Language Therapist and Educational Support Assistants

The first aspect to be examined was how effective the stakeholders considered the technology to be. ESAs and the SLT came from a range of technical backgrounds which did not seem to affect their responses to the questions asked of them. Their computer experience and assessment of the technology is outlined below. The numbers of respondents are in brackets.

- 50% (3) stated they were 'quite experienced with computers', 33% (2) were 'moderately experienced with computers', and 17% (1) had 'no experience with computers'.
- 42% (5) of ESAs / SLT stated that it was 'very easy' to hear the remote person, while 58% (7) stated that it was 'fairly easy' to hear the remote person.
- 67% (8) stated it was 'very easy' and 33% (4) stated that it was 'fairly easy' to ask questions across the link.
- 17% (2) were unsure whether an audio delay occurred, while the remaining 83% (10) reported an audio delay.

- When asked to compare face-to-face with video-mediated communication based on how easy it was to explain their point, 58% (7) answered 'very easy', 25% (3) answered 'fairly easy' and 17% (2) answered 'neither easy nor difficult'.

- When considering the video image, 16.6% (2) of ESAs / SLT rated the image as 'excellent', 66.6% (8) rated it as 'very good' and 16.6% (2) rated the image as 'average'.

- 83% (10) of ESAs / SLT reported that they looked at the remote person 'very frequently' and 17% (2) reported it as 'fairly frequently'.

- 100% (11 - one respondent did not answer this question) of ESAs / SLT felt that the video image helped them put their point across and helped them understand what the remote person was saying.

Social aspects of the use of the technology were also investigated. These include satisfaction levels, whether stakeholders felt able to chat openly across the videoconferencing link and whether the stakeholders had met the remote collaborator before sessions started across the link.

- 67% (4) of stakeholders had met the remote person before the videoconferencing sessions had started. This question was not included in the follow-up questionnaire.

- 100% (12) of ESAs / SLT felt able to focus on the agenda of the session, understand and take in what the remote person was saying.

- 100% (12) felt they could chat informally with the remote person.

- When completing the questionnaire for the first time, 71% (5) felt 'very satisfied' and 29% (2) felt 'fairly satisfied' with the outcome of the sessions and from completing the follow-up questionnaire 80% (4) felt 'very satisfied' and 20% (1) felt neither satisfied nor dissatisfied with the outcome of the sessions.

The follow-up questionnaire included a section for respondents to compare their present and past experiences.

- 100% (5) of respondents indicated that they felt their communication had changed since they had started using the videoconferencing system. All re-

spondents indicated that this change was positive. Example responses indicated 'it feels more like face-to-face communication now' and 'talking into the microphone is easier'.

- 40% (2) of ESA respondents indicated that their working practice had changed as a result of the videoconferences with the SLTs. When asked to explain their answer, reasons given included 'I feel more experienced in my line of work' and 'there are more therapeutic techniques available for me to use'.

Parents

Questionnaires explored how the parents viewed the care their child was receiving, how they felt about the technology being used to deliver support and investigated their computer experience.

Satisfaction levels

- 43% (3) of parents were 'very satisfied' and 43% (3) were 'fairly satisfied' with the teaching their child was being given by the ESA. Results from the follow-up questionnaire indicate that 100% (5) of parents were 'very satisfied' with the teaching standards of the ESAs.

- 75% (6) of parents were 'very satisfied' and 12.5% (1) were 'fairly satisfied' with the progress of their child. Results from the follow-up questionnaire indicate that 80% (4) of parents were 'very satisfied' and 20% (1) were 'fairly satisfied' with the progress of their child.

Parents' views of the technology are outlined below. See Appendix I for the full results.

- From both the first and second questionnaires, 100% (13) of parents had no reservations regarding the SLT using the link.
  - from both questionnaires, 77% (10) of parents could see advantages in using the link.
  - several parents also commented on the advantages that could occur if the ESA and the SLT saw each other face-to-face.

- 50% (6) of parents could see disadvantages with using the technology; an example given was increased expense.
- Parents came from a range of technical backgrounds that did not seem to relate to their questionnaire responses.

3.4.2 Comparison Group - Questionnaires

The first sets of questionnaires were sent to the comparison SLTs in December 1997. At the beginning of January 1999, questionnaires were sent to the comparison parents. Response rates are indicated in Table 3.2. The material from the questionnaires is presented below. Data which are deemed to be of particular interest are shown in this section and the full results of these questionnaires are given in Appendix H (SLTs) and Appendix I (parents).

Speech and Language Therapist

Questionnaires sent to the SLTs examined their working habits and environment. Results indicated that the comparison SLTs had vastly differing working environments from each other and from that of the TeachSpeech SLT; for example, case loads and amount of contact with ESA varied considerably. These differences are reflected in their working practices.

Parents

The questionnaires investigated issues about the type of therapy their child was receiving, how satisfied they were with the therapy and how they would feel if their child was receiving therapy using the TeachSpeech model of therapy.

- 33% (3) of parents were 'very satisfied', 44% (4) were 'fairly satisfied' and 11% (1) were 'neither satisfied nor dissatisfied' with the support their child was receiving. One respondent (11%) indicated that they were 'very dissatisfied' with the support; the reason given was that they felt their child was not receiving any support.
- 44% (4) of parents were 'very satisfied', 33% (3) were 'fairly satisfied' and 11% (1) were 'neither satisfied nor dissatisfied' with the progress of their child.

Table 3.2: Comparison group questionnaire completion rates where R represents a reminder questionnaire which was sent to the parents a few weeks after the original questionnaires were sent

Category	Category Population	Number of Respondents	Percentage of Respondents
SLT	2	2	100
Parents	16	4	25
Parents R	16	$5 + 4 = 9$	56

Table 3.3: Direct comparison of the satisfaction levels of parents, across both service delivery models, regarding the progress of their children

Category	TeachSpeech Group	Comparison Group
Very Satisfied	75% (6)	44% (4)
Satisfied	12.5% (1)	33% (3)
Neither satisfied nor dissatisfied	/	11% (1)
Dissatisfied	/	/
Very dissatisfied	/	11% (1)

Again one respondent (11%) indicated that they were 'very dissatisfied' with the progress of their child since no support was being delivered.

- 78 % (7) of parents indicated that they would have no reservation if the SLT used a videoconferencing system to give support to the ESA.
- 89 % (8) indicated that they could see advantages in using such a system.
- When asked if they thought that the quality of care would decline if the SLT supported the ESA either face-to-face or over the link, 56% (5) replied no.

Since the questionnaires given to the TeachSpeech parents were understandably different from those given to the comparison group, only one direct comparison can be made. This direct comparison is in relation to the satisfaction levels with the progress of the children. Overall, both sets of parents seemed satisfied. However, it does seem that the TeachSpeech group was marginally more satisfied with the service that was being provided as shown in Table 3.3. From interviews it was established that the one parent who was dissatisfied, was dissatisfied with the amount rather than the quality of therapy her child was receiving.

Table 3.4: Illustrates when and with whom the interviews were conducted

Date	SLT	ESA	Parents	Teachers	Head Teacher
March 1998	1	4	2	1	2
November 1998	0	0	5	0	0
May 1999	2	4	5	1	2

3.4.3 Interviews

All stakeholders who had an interest in the TeachSpeech project were invited for interview. The semi-structured interviews were tailored according to the role of the individual. The interviews were held three times during the evaluation period as seen in Table 3.4.

The interview data indicated that both the SLTs and ESAs had a positive attitude towards the TeachSpeech model of therapy. The interviews with the TeachSpeech SLT highlighted her ability to change treatment goals when necessary, being alerted quickly to any new speech or language problem a child was having, and being able to provide support to the ESA when unusual problems occurred.

Other particular points of interest include a comparison of the two models, confidence levels, and usability of the videoconferencing system. The majority of the responses indicated that the stakeholders believed that the TeachSpeech model of therapy was an effective and efficient method for delivering speech and language therapy. Generally, the ESAs indicated that their confidence levels had increased especially in the support they provided children in the one-to-one situation. There were a few usability problems, including the need to replace desktop microphones with tie microphones.

Four ESA were interviewed twice, the first time in March 1998 and the second time in May 1999. It is worth noting that their opinions of the TeachSpeech model of therapy remained positive.

Although parents were often sceptical at first towards the TeachSpeech model, they stated that with time they became more positive. One parent who was interviewed twice was extremely positive about the TeachSpeech model. She believed that her child had progressed well and that having the link within the school had many other positive consequences.

3.4.4 Comparison Group - Interviews

The semi-structured interview investigated how the SLT felt about the TeachSpeech model of therapy. The advantages included the time saved travelling and having access to more children. When asked about any disadvantage the TeachSpeech model of therapy might have, the opinion that nothing could replace face-to-face therapy between a child and a SLT was raised.

The responses from the parents in the comparison group varied and seemed to fall into two categories. Either a child was receiving therapy regularly, or they were not receiving therapy at all. The former group generally deemed quality of care as high, while the latter group seemed frustrated that no therapy was being given. When asked to compare the two models of therapy, comments again varied widely. Several comments were highly positive regarding the prospect of the utilisation of a different model of therapy, while other comments stated reservations regarding the use of a videoconferencing system.

3.4.5 Demonstration Programme Questionnaires

The demonstration programme questionnaires were designed to look at how various professionals, who may in the future have the potential to use such technology as part of their working practice, felt about the videoconferencing technology. They were asked to assess whether they felt such a system would be valuable for either their profession or another profession. This is a useful part of the evaluation as it plays a part in assessing how a wide variety of professionals initially viewed the technology.

Those who attended the demonstrations included clinical psychologists, teachers and SLTs. The questionnaires investigated how valuable the technology might be from the perspective of the respondents' professions, and from the perspective of other professions. It inquired as to how useful the technology would be in enabling the respondent to give input to others or to receive input themselves. These questions were asked before and after a demonstration of the technology was given.

There were four patterns of responses and the majority fell into the first type. Firstly, those that liked the technology before they saw it being demonstrated liked it after the presentation. There were 55 of these and they responded positively all the way through the questionnaire. The second group comprised of 3 respondents, who

did not say whether other professions would find the technology useful. After the presentations, 2 of these respondents indicated that they thought other professions would find the technology useful while the remaining respondent was uncertain. The third group consisted of 2 respondents who replied 'don't know' when asked whether the link would be useful to provide input to others and themselves. After the presentation both the respondents changed their answer to indicate that they thought the system would be useful. Finally, 4 respondents answered 'don't know' to all questions concerned with the usefulness of the technology. Of the 4, one answered 'no' as to the question of usefulness to providing input to others. After the presentation, all respondents changed all their answers to 'yes', except for the respondent mentioned above who changed their answer to 'don't know'. It should be noted that one respondent did not fit into any of these categories. This respondent answered 'don't know' both to whether the technology had future value for other professions and to whether the technology would be useful to provide input to them. Both answers were changed to positive after the presentation.

In summary, the vast majority of the respondents answered positively to the questions posed. For full details, see Appendix J.

3.5 Discussion

From the results obtained from the questionnaires and interviews, it seems that both models of therapy were deemed to be of high quality by stakeholders, namely the parents, the teachers, the ESAs and the SLTs. To reiterate, the hypothesis is,

- stakeholders will be equally satisfied with both models of therapy.

3.5.1 TeachSpeech Group

If we combine the results from the first and second questionnaires, 75% of the Teach-Speech ESAs and SLT felt that the outcome of the sessions was 'very satisfactory' and 17% rated the outcome of the sessions as 'fairly satisfactory'. This is consistent with past research into patient satisfaction in telemedical consultations (Callahan *et al.*, 1998) in that it seems that there is a general trend which demonstrates high satisfaction levels for telemedical videoconferencing sessions. These data are highly encouraging, and point towards a high quality of interaction taking place. During

the evaluation interviews, several comments were made highlighting how focused sessions were, with stakeholders feeling they were able to focus extremely well on the topic being discussed. This may explain in part the high levels of satisfaction reported.

The results from questionnaire data focusing on the effectiveness of the technology were found to be highly positive. Stakeholders felt that the audio and visual quality of the videoconferencing system was high. The videoconferencing system was not only being used for verbally reporting information but other productive activities also occurred. For example the ESAs displayed work the children had done to the SLT and the SLT displayed books the ESA could use as part of therapeutic exercises. From comments made on the questionnaires, it was evident that visual aids could be seen clearly when they were held near the screen.

Interviews also provided data on stakeholders' perceptions of the quality of the therapy. ESAs were asked to compare directly the two models of therapy and all answers either stated that the two models were comparable, or that there was an advantage for the TeachSpeech model. In addition, the follow-up interviews and questionnaires tended to confirm initial positive responses. The time or amount of use of the videoconferencing system did not seem to have affected attitudes.

3.5.2 Comparison Group

From the interview and questionnaire data it was found that parents were generally satisfied with the therapy their child was receiving. Those that exhibited dissatisfaction generally did so because their child was receiving no therapy, rather than perceptions of the quality of therapy being poor. Overall, satisfaction levels were high.

Questionnaires were used to investigate attitudes towards the possible future developments of the TeachSpeech model of therapy. Parents were asked to envisage their child receiving therapy from the TeachSpeech model. Seventy-eight percent of parents indicated that they would have no reservations about their children receiving support from a SLT via a videoconferencing link.

3.5.3 Demonstration Programme Questionnaires

This questionnaire examined how professionals from a wide range of fields, from clinical psychologists to occupational therapists, felt regarding the videoconferencing technology and whether they thought that this technology would be useful to their own or another profession. The majority of responses in both these areas were positive. Furthermore, those who either answered negatively or who were unsure before the presentation tended to change their responses after the presentation had been given. Sixty out of the sixty-two questionnaires answered positively to the question as to whether they thought the technology would be useful to them in their working environments.

3.6 Conclusions

It is clear from the results presented that the video technology is accepted and liked. This is supported by evidence from the questionnaires and the semi-structured interviews. This finding is consistent with previous research. For example, Johansson *et al.* (1997) conducted a study which involved speech pathologists communicating with their patients using videotelephony. Results showed that both users and their partners accepted the medium. Zarate *et al.* (1997) used videoconferencing to assess the severity of symptoms in patients with schizophrenic illness. They found that the technology was effective in obtaining an evaluation of the patient and that users accepted the technology. Finally, Callahan *et al.* (1998) demonstrated that 66% of non-mental health patients using a telemedical system perceived the method as better than face-to-face care. For the TeachSpeech project, these findings hold both for those who were using the technology and prospective users of the technology. It can be seen that the evaluation of stakeholders' perceptions has added a great deal to the evaluation overall. The results of this analysis have added a unique dimension to the evaluation of the TeachSpeech project as the perceptions of users are an important component in assessing the overall success or failure of a project and a study which lacks such data is surely missing an important perspective.

Chapter 4

Evaluating Process

4.1 Introduction

This chapter aims to provide a description of the process of mediated communication that occurs between speakers in the TeachSpeech model of service delivery. Due to considerations discussed in Section 1.1.3, only dialogues from the TeachSpeech group are recorded and discussed here. This description serves to assess and appraise the communication. From the previous review in Section 1.4, it was decided that structurally analysing and coding the dialogues would be informative.

Analysing the process of communication will be a valuable addition to the evaluation of the TeachSpeech project for the following four reasons; firstly, as outlined in Section 1.4, mediating an interaction can affect the communication process. It would therefore appear important to describe and appraise the communication process. Secondly, it is hypothesised that evaluating the process of communication will allow for further insight into the findings attained in the evaluation of performance and perceptions. Thirdly, as a research method it has the benefit of being unobtrusive. Consequently, the data gathered are naturalistic and there is little danger that the measurement tool has affected the data. Finally, in the past it has proved a useful and effective technique for providing guidance as to the type of interaction occurring, for example, in terms of interactivity (Sellen, 1995) or efficiency (Boyle *et al.*, 1994).

The aims of analysing the process of communication were twofold. The first aim was to provide a description, in terms of content and structure, of the communication that occurred between speakers. Specifically, we wished to investigate the

listribution of communication topics raised by the speakers and how these changed over time. For example, whether social interaction or communication concerning the technology occurred. We also wished to assess whether the quality of the present system is sufficient to sustain complex interaction such as the process of teaching and the extent to which the visual channel was used.

4.2 Overview of Chapter

The following section examines a variety of content coding schemes, and discusses the relative merits and drawbacks of each in turn as applied to the aims of the current analysis. Based upon this review, an overview of the coding scheme used in the present analysis is presented. Section 4.7 addresses the first aim of this chapter, namely, to determine what topics are being discussed during the video-mediated meetings between the SLT and ESA. This is done by providing results from the content coding and structural analysis. The following section provides a discussion of these findings. Section 4.7.2 investigates the second aim of this chapter which is to determine the value of the visual channel. Results and discussion are provided. Finally, conclusions are given in Section 4.9.

4.3 Coding Schemes

Ways of analysing discourse have originated in a diverse number of fields including psychology, philosophy, sociology and linguistics. While some coding schemes such as Conversational Analysis (Schegloff *et al.*, 1977; Sacks *et al.*, 1974) analyse, for example, the way in which turns of talk are passed between speakers, others have looked at the content of dialogues. The aims of the coding scheme for this research have been outlined above and, in brief, include providing a description and evaluation of the process of communication that occurs between speakers across the videoconferencing link in the TeachSpeech model of service delivery. Thus, it is the content of the dialogues which is of primary interest. Three such content coding schemes will be outlined here, with the intention of evaluating these schemes with respect to the aims of the current analysis.

The coding schemes which will be evaluated here include Bales' (1950) Interaction Process Analysis (IPA), Olson *et al.*'s (1992) scheme for coding design meeting

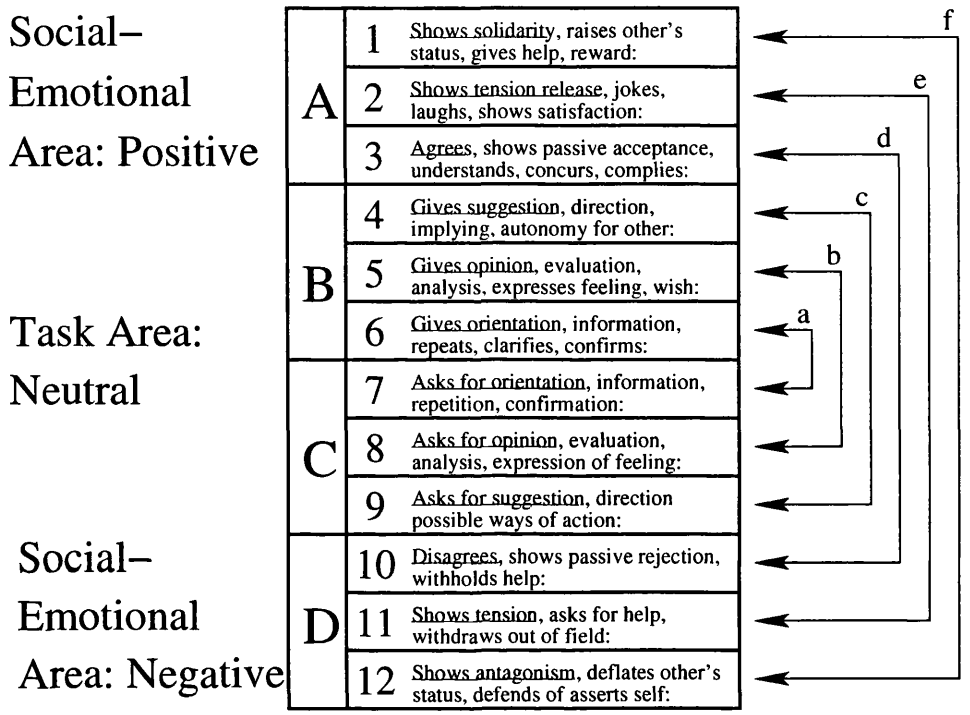
activity and the Conversational Games Analysis (CGA) developed by Kowtko *et al.* (1991). The usefulness of each scheme in relation to the aims of this analysis will be highlighted and research which has utilised the respective schemes will be described. Conclusions are drawn concerning whether these coding schemes would be useful for the research here. The usefulness of each scheme for the TeachSpeech project in either a raw or a re-defined version will be considered, as well as the original aim of the coding scheme and its reported reliability.

4.3.1 Interaction Process Analysis

Bales (1950) developed Interaction Process Analysis (IPA) as a theoretical framework in which to code small group interaction. This scheme is applied to speech from both the addressee and addresser. The unit that is coded is defined as the 'smallest discriminable segment' of speech. Every observed act of verbal and non-verbal communication between participants is allocated into one of twelve categories. These twelve categories in turn fall into four main areas, as illustrated in Figure 4.1.

The main aim of Bales' (1950) IPA system was to enable a coder to categorise entire dialogues. The scheme was intended to be sufficiently flexible to code dialogues in which different topics were discussed and in which the number of participants differed. The reliability of the IPA was reported in a study by Heinicke and Bales (1953), where a Pearsons r of 0.86 for two coders was given. Waxler and Mishler (1966) reported a kappa of 0.61 (n=10910, κ=2). Bales notes that much practice and frequent training is required to attain good reliability.

Using this scheme, Bales (1955) examined which type of interaction occurred during a face-to-face group discussion task. Bales asked groups of two or three participants to complete a complex task, where the groups were leaderless and unfamiliar. Results showed that 56% of acts were coded as problem solving attempts and the remaining 44% were distributed between the other categories including positive reactions, negative reactions and questions. Within a meeting, a pattern emerged as to the types of acts that were uttered. 'Giving information' tended to be most frequent in the first third of the meeting, and then declined for the remainder of the meeting. 'Opinion giving' was more frequent in the middle of the meeting and 'suggestions' were highest in the last third of the meeting. Rates of positive and negative reactions were at their lowest at the start of the meeting and they steadily



KEY

A Positive reactions
B Attempted answers
C Questions
D Negative reactions

a Problems of communication
b Problems of evaluation
c Problems of control
d Problems of decision
e Problems of tension reduction
f Problems of reintegration

Figure 4.1: The system of categories in Bales' Interaction Process Analysis

increased throughout the ensuing stages of the meeting. The author suggests that once a decision was made, negative reactions decreased, and at the same time positive reactions increased. Bales (1955) also found that the best-liked individuals were usually placed second or third in the participation hierarchy.

Bales and Borgatta (1955) conducted further analysis using the IPA system. They investigated effects of group size and the variability of individuals' communication style over given sessions. Results suggested that when the group size increases, the frequency of utterances coded as 'tension release' and 'giving suggestions' also increase, whereas 'showing tension', 'showing agreement' and 'asking for opinions' decrease as groups size increases. The authors suggest that these trends are a result of two factors. Firstly, the amount of time each person has to talk decreases, and secondly each person has more people to talk to as group size increases. Another effect of group size was that individual profiles were affected. These seem to have been affected mostly by whether the group consisted of an odd or even number of people. The researchers suggest that this was due to more conflict occurring in even groups since the possibility exists for deadlock. In groups of even numbers, profiles of individuals tended to be similar, that is, variability was low within categories, with the exception of giving suggestions. The authors report that, as group size increases, there is an increased variability in the amount of suggestions given. This seems to indicate that as group size increases, the probability of an obvious leader emerging declines. They also found that individuals who have a high variability among categories tended to have low interaction rates.

Bales' IPA can be criticised for several reasons. On a general level, Bakeman and Gottman (1997) recommend that coding schemes should be developed with the aim of answering a particular hypothesis. Bales' IPA is a general coding scheme which was intended for application in many different small group interactions. Morley and Stephenson (1997) also noted that various categories in the IPA system are not mutually exclusive which is an undesirable trait in coding schemes. With respect to the current research, other drawbacks of this coding scheme can be highlighted as follows. Firstly, Bales' IPA was designed to investigate broad differences between small groups whose discussion topics ranged widely. The current research intends to investigate two person interactions, with the aim of exploring the fine detail of such interactions. Secondly, Bales' IPA originated from his observations of an Alcoholics

Anonymous group (Bales, 1955) and was designed to investigate social and cultural aspects of group interaction. These original intentions differ greatly from the intentions of the current study. Finally, Waxler and Mishler (1966) reported that the reliability between coders was not good, as they only managed to obtain a kappa of 0.61. Taking these factors into consideration, Bales' IPA does not seem suitable for use with the current study.

4.3.2 Design Meeting Activity Coding

Olson *et al.* (1997) developed a scheme for coding meeting activity. The categories they used were inspired by categories of group management processes described by Putman (1981) and Poole and Hirokawa (1986). The scheme aimed to examine participants' problem solving strategies and the activities they used to co-ordinate and manage themselves. There are two levels to this scheme, the first of which has 11 categories. Table 4.1 gives a brief definition of each category. After a preliminary analysis of the dialogues, the researchers found that 'clarification' was a major category, thus a second level of analysis was developed. Every category was then split into two, its original form and a clarification form of that category. For example, the category 'alternative' became 'alternative' and 'clarification of alternative'. Two other specific categories also devised were clarification of artefact and clarification general. Clarification artefact refers to a participant explaining what is meant by a list, object or artefact. Clarification general refers to any clarification that could not be classified in any of the other categories. Reliability for the scheme was assessed by Olson *et al.* (1992). Three coders produced inter-observer reliability correlation between 0.68 and 0.99, with a median of 0.85.

Olson *et al.* (1992) explored both the amount of time speakers spent on each topic and the transitions that occurred between topics. Their findings showed that although the numbers of issues raised differed in the ten design meetings they coded, the distribution of topics was remarkably similar. It was reported that 40% of the time was spent on discussion of the design, while approximately 30% of the time was spent on summaries and walkthroughs. The authors also noted the extensive use of clarification time and suggested that this related to a drive for common understanding, or as Clark and Brennan (1991) term it, 'common ground'.

Olson *et al.* (1997) investigated meetings that were held across different

Table 4.1: A brief description of first level categories of Olson *et al.*'s (1992) design meeting coding scheme

Category	Definition
Issue	The major aspects or problems of the designed object itself. This would include any elaboration of these ideas.
Alternative	Proposals or suggestions regarding aspects of the design feature. Usually involving which features to offer and how to implement them.
Criterion	Evaluation of an alternative solution or proposal. This includes reasons, arguments and opinions that appraise a proposal, sometimes involving an analogous system.
Project Management	Statements concerning the assignment of tasks which are not directly related to the design under discussion. Deciding when to meet again.
Meeting Management	Organisation of the meeting itself, agenda management.
Summary	Review of the state of the design. Simple list like statement giving overview of previously discussed topic.
Clarification	Questions and answers which clear up misunderstandings.
Digression	Covers comments which are not topic related, jokes and any utterance from camera operator.
Goal	Statements regarding the main aim of the meeting. Motivational statements.
Walkthrough	An overview of the design so far. Usually outlining the procedure the user will undertake.
Other	A category to encompass all other comments.

communication media using different computer support tools. This is particularly relevant to the TeachSpeech project and thus a brief overview is given. The four conditions examined included: a face-to-face meeting using a whiteboard and paper and pencil, a face-to-face meeting with a shared editor, a remote meeting with high quality audio link and the shared editor, and a remote meeting with high quality audio and video links with the shared editor. All sessions were transcribed and coded according to the above scheme, although four alterations were made. Olson *et al.* (1997) included the categories plan, write, technology confusion and technology management. Plan refers to occasions when the participants would plan the organisation of a meeting while write refers to the participants discussing wording or dictating during the meeting. Technology confusion encompassed topics where the participants were having difficulty with the technology, and technology management referred to the placement of work onto the computer interface. As in their previous work, they found that the groups spent their time in a similar manner. Nevertheless, they did find that remote groups spent more time clarifying and managing their meeting, although the video condition spent less time clarifying issues as compared to the audio condition. This coding scheme therefore illustrates a benefit in utilising the visual channel.

On a general note, the coding scheme is highly specific. In the second study described above, the researchers changed the design of their study slightly and in doing so four new categories needed to be added. The main limitation as related to the TeachSpeech project pertains to the fact that the coding scheme was originally devised for use within design meetings and this is reflected in the majority of the categories. This relates to the above limitation of specificity. If this scheme were to be used to analyse the TeachSpeech dialogues, it is likely that the coding scheme would need to be extensively altered. Secondly, the aims of the current analysis rely on a coding scheme which will reveal the nature of the interaction taking place. For example, Olson *et al.* (1992) may code an utterance as alternative, however as part of the aims of the current analysis there is a need to comprehend the content of the alternative utterance as it could relate to a speech and language therapy technique or how to best describe a particular speech and language impairment. Olson *et al.*'s (1992) coding scheme assumes that all issues discussed are of the same type, that is, design problems. Within the TeachSpeech dialogues, the SLT and ESA encounter

a wide variety of problems pertaining to the videoconferencing link and this would not necessarily be captured by this coding scheme. In summary, although Olson *et al.*'s (1992) scheme achieves adequate inter-rater reliability, there are limitations to the scheme in regard to coding the TeachSpeech dialogues and therefore the scheme seems inappropriate.

4.3.3 Conversational Games Analysis

Kowtko *et al.* (1991) developed the CGA coding scheme, which codes the functional use of utterances in task-oriented dialogues. It is based on work from the field of artificial intelligence (Houghton and Isard, 1987; Power, 1979). The use of this coding scheme involves coding every utterance according to what the speaker is attempting to achieve, by the function of that utterance. Carletta *et al.* (1996) describe the three levels of dialogue structure that can be analysed using this system. The scheme segments the dialogues into transactions, conversational games and conversational moves, with each level becoming more detailed. The first level, 'transactions', are sections of the dialogue which allow participants to draw a segment of a route, as part of a collaborative problem-solving task. These are manageable sections. The second level of analysis, is termed 'conversational games'. These are sections of the dialogue where an initiation occurs which either seeks information from another participant or provides information to another participant. A conversational game ends when it has either been fulfilled or abandoned. The final level is termed 'conversational moves'. Just as the conversational games make up the transactions, the conversational moves make up the conversational games. Carletta *et al.* (1996) describe conversational games as various kinds of initiations and responses which are classified according to their purposes.

Calculations of reliability have been good. Two coders applying the scheme to the map task, in which two participants are required to navigate around a map via verbal instructions, produced a kappa of 0.7 (Anderson *et al.*, 1997). The scheme has also been applied to a range of two person face-to-face and technologically mediated interactions, receiving kappas of 0.95 (Carletta *et al.*, 1997) and 0.94 (Newlands, 1998), respectively.

Research using CGA has explored several areas of communication. Doherty-Sneddon *et al.* (1997) applied CGA to a selection of problem solving dialogues, from

the Human Communication Research Centre (HCRC) map task corpus. Doherty-Sneddon *et al.* (1997) explored spoken dialogues where speakers were visible or screened from their listeners. Although performance was unaffected by the communication medium, the dialogues were significantly longer in the audio-only context as the participants were using more align and check games; align games provide feedback to the listener and check games involve checking your own understanding of a previous utterance. The authors suggest that the participants in the audio-only condition were ensuring mutual understanding was taking place verbally, whereas in the face-to-face condition participants were using non-verbal signals.

More recently, Veinott *et al.* (1999) used a simplified form of CGA to assess how native speakers versus non-native speakers of English proceeded to complete a variation of the map task when supported by different kinds of media. Their study involved native and non-native speakers completing the map task with either video and audio, or audio-only conferencing. They believed that the non-native speakers, whose background and thus common ground differed, would benefit from the visual channel more than the native speakers. The non-native speakers would be able to understand the other participant more clearly with the aid of video and could monitor the understanding of their collaborator visually. Overall, native speakers obtained higher performance scores although, as expected, the non-native speakers' scores increased when video was introduced. The dialogue analysis revealed that the non-native speakers spent significantly more of their time clarifying their contributions. The authors suggest that the overall performance loss for the non-native speakers was associated with less instructing and checking of their mutual understanding, which resulted in less grounding occurring in the audio-only condition.

CGA was intended to be able to represent dialogues from a variety of origins and it has proved extremely useful when applied to dialogues which result from tasks such as the map task and other structured problem solving tasks. It does, however, contain several limitations in relation to the current study. Firstly, problems can emerge when applying this coding system to naturally occurring dialogue. Usually, natural dialogue is not orderly and has no predetermined goals and, if such goals do exist, they are not explicitly stated and as such, CGA may be less useful for exploring unstructured interaction such as the dialogues obtained from the current research. Secondly, with CGA the main emphasis lies with how information is shared rather

than what topic of information is introduced. With the TeachSpeech project we are interested in the latter function, namely, communicative topic. Clearly, these are two different dimensions by which to code dialogues. While the reliability for CGA is excellent and is an important consideration, the main purpose of the present analysis concerns the communicative topic of the speaker and thus coding the dialogues using CGA would be inappropriate.

4.3.4 Summary of Coding Schemes

There are a variety of different ways to code dialogues. The three presented here show how content and functional analysis improve our understanding of communication. Bales (1950) demonstrated the range of communication functions that occur in group meetings and found the patterns which emerge are relatively stable over time. Bales and Borgatta (1955) investigated the impact of group size and individual profiles on group discussions. They showed that individuals communicating within an even numbered group as compared to an odd numbered group, tend to have more stable profiles over time . Olson *et al.* (1992) coded the communication of participants who took part in design meetings. They showed that the amount of communication that occurred in each topic was remarkably similar over different conditions, but that clarification occurred more frequently in mediated communication. Anderson *et al.* (1997) outlined the effects on communication of adding a visual channel by using CGA, from which they found that there were more align games in audio-only communication, indicating that speakers needed to check that listeners had understood what they were saying more frequently.

As described, the coding of dialogues is a useful additional research tool which can reveal detailed aspects of communication. It can provide evidence about the nature of the interaction which may run contrary to self-report. It is well known that introspection may be an unreliable way of collecting data on performance and therefore it is necessary to go beyond questionnaires and interviews for an in-depth investigation of the impact of technology on communication. As indicated, however, the previously described coding schemes are not suitable for the present aims, namely to

- describe the communication that is taking place, and to
- examine the value of the visual channel.

In order to meet these specific goals of the current work, it was necessary to develop a coding scheme which would address the above requirements. The following section overviews essential considerations which must be addressed in order to develop a new coding scheme, and applies these criteria to a coding scheme in order to satisfy the present aims.

4.4 Development of a Coding Scheme

Bakeman and Gottman (1997) stipulate that a researcher should never use a coding scheme developed by another person. They suggest that a valuable scheme is one which is designed with a specific hypothesis in mind. Methodological design issues have been investigated by Bakeman and Gottman (1997), who provide several guidelines for developing a coding system. These include: beginning with a clear question, keeping the scheme simple, ensuring the detail of analysis within each category is consistent, developing unique and distinct categories and, of critical importance, reliability. Each of these will be considered in turn.

There are many aspects to take into account when developing a coding scheme and one of the most significant relates to what question the researcher is asking. It is important to begin with a clear question. Since developing a coding scheme is theoretical, it should fit the specific ideas, questions and hypotheses of the researcher. The development of a coding scheme which captures the content of a dialogue is usually highly specific to a particular research question. It has been suggested that the success of observational studies often depends on the development of a thoughtfully constructed and well-formed coding system. If the coding system is well organised, it should paint a clear picture of the major components relevant to the research question.

It is also sensible to keep the coding scheme simple. Bakeman and Gottman (1997) believe that this will help ensure the other important components of the scheme are upheld. Many reasons for this were given. Firstly, if the categories are clear and are essentially at the same level of description, the coders will be able to code the transcriptions more reliably. Secondly, codes should be mutually exclusive and exhaustive. Coding schemes which are said to be mutually exclusive follow the premise that each unit of categorisation is classified into one category only. As mentioned previously, Bales' (1950) IPA was susceptible to this criticism.

Exhaustiveness refers to the scheme having categories that capture the entire dialogue. While it is not a requirement that coding schemes are mutually exclusive and exhaustive, such schemes do exhibit certain beneficial characteristics. To construct this type of coding scheme requires an amount of conceptual analysis and their use can simplify data analysis.

Reliability of a coding scheme is essential and the two main types are inter-judge reliability and reliability decay. Each will be examined in turn. It is necessary to examine inter-judge reliability since a particular coder may have a unique perspective on how the transcripts should be coded and thus it is a requirement to train at least two judges and ensure agreement occurs between them (Bakeman and Gottman, 1997). However, it has been noted that this only addresses potential errors between judges and ignores other sources of error. It is possible that the two judges trained may have their own unique perspective, so further measures need to be undertaken to increase reliability. A coding scheme needs not only to ensure that the judges agree with each other but also that this agreement does not diminish over time from a process referred to as reliability decay. Taplin and Reid (1973) have demonstrated that reliability decay does indeed exist. They trained participants to code to an accuracy level of 80% and found that over time the accuracy level decreased. This can be critical if either the coding system is complex, or the actual process takes a long time. It is therefore necessary to ensure the transcripts are being coded consistently over time.

4.5 The Current Coding Scheme

The overall aim of the current coding scheme is to describe and investigate the communication that takes place between the SLT and ESA over the videoconferencing link, and to appraise the use of the visual channel

A preliminary coding scheme was developed which included seven categories: administration, evaluation, reporting, social, teaching, technical and other. These categories were devised by a combination of quantitative and qualitative research methods, for instance, by reviewing relevant literature and examining the transcripts in detail. It has been argued within a realist framework that coding categories can be discovered from within the data (Madill *et al.*, 2000). Certain categories have been devised using this method. A coding scheme which was used as a general guideline

for the present analysis was the scheme used by Carletta *et al.* (1998). They coded communication in two virtual supply chain teams who were using desktop video conferencing for collaborative work. Their categories included utterances which were related to the technology, utterances which were social, those which directly addressed the task which concerned an information source (e.g. a CAD diagram), and lastly those which concerned an external distraction.

The seven categories that were devised as part of the current coding scheme are presented in Table 4.2. These categories can be divided into two main components: non-task talk and task talk. The non-task talk categories will be described first; these include talk concerning the technology, social talk and evaluation talk. Secondly, task talk categories consist of teaching talk, reporting talk and administration talk.

The technology category included any speech which was related to the videoconferencing system. Usability issues which concerned the audio and video capabilities of the videoconferencing system, comments regarding the microphones, and the position of the remote participant are all included in this category. Carletta *et al.* (1998), Olson and Olson (1992), and Wright and Monk (1989) all included similar versions of this category in their coding schemes. Carletta *et al.* (1998) categorised utterances in this category if they concerned the usability of the technology. Olson and Olson (1992) included two categories termed technology confusion and technology management. The former related to the confusion regarding the technology while the latter concerned the placing of work onto the interface. Wright and Monk (1989) coded all instances where the tools available to the participants were mentioned directly in the dialogues. They termed this category as breakdowns. In their research, only a few such instances occurred but their work was laboratory based. The TeachSpeech project is a field study in which the SLT and ESAs operated the videoconferencing system themselves with only minimal training. After careful consideration, it was concluded that the TeachSpeech dialogues required a technology category. The inclusion of this category was mainly based on previous research in this area.

The social talk category included all talk that was purely social in nature. Greetings, farewells and casual conversation regarding social aspects of the participants' lives were allocated to this category. Carletta *et al.* (1998) coded their

Table 4.2: Categories utilised in the current coding scheme. It should be noted that all names have been changed, in this table, and in all other instances.

Category	Definition	Example
Technology	Any speech regarding the video conferencing system including the video image, the audio quality or the microphone. This category is exclusively for the videoconferencing system and not for any other equipment.	ESA: Oh oh hang on you're a bit quiet
Social	Social speech such as greetings, informal chat and good byes.	SLT: Ok take care ESA: See you SLT: Bye Jane
Evaluation	Any speech concerning the evaluation of the project. Any speech regarding the interviews, questionnaires, the evaluator's visits to Wiltshire or the audio recording of the sessions. Also any speech spoken by the evaluator.	SLT: Yeah ok you know I'm recording you know I'm recording all of our sessions now ESA: Yeah SLT: Ok just so you know
Teaching	Speech where the SLT is instructing or giving tuition to the ESA. This may involve teaching a new technique to elicit a particular sound or word.	SLT: Yeah when you you start to work on initial S it's sometimes better to start on blends rather than the initial S because you can neatly separate it from the from the second sound so S pot and S in snowman
Reporting	Any speech relating to the progress of the child or the child's family regarding speech and language therapy	ESA: All right can we kick off with Mary SLT: Mary yes ESA: yes ok we've seen a full event on all P sounding words, that we were having a problem with
Administration	Any speech concerning the arrangement of meetings either in face-to-face, video-mediated or telephone scenarios. This category also included any speech regarding speech and language therapy that is not directly related to the task at hand.	ESA: Am I speaking to you next week? SLT: Yes ESA: Yes SLT: Yeah yeah I'll speak to you next week

dialogues for social information. They suggested that social interaction is important in building long term working relationships and they felt this was an important category to include. From the TeachSpeech dialogues, it was noted that social interaction regularly occurred during the initial and final stages of the meetings and thus it seemed wise to include this category.

The final category in the non-task talk section was evaluation talk, which includes any speech regarding the evaluation of the TeachSpeech project, and includes, for example, communication concerning questionnaires or interview dates. It was felt that talk concerning the evaluation of the TeachSpeech project needed to be distinguished from the other categories. Evaluation talk would not usually be part of a session between and SLT and an ESA and therefore it would be useful to classify this category separately. The evaluation talk category is clearly unique to the TeachSpeech dialogues. If this category did not exist, this type of talk would be allocated to the other categories which is undesirable and therefore it was decided that a category for this would be necessary. This category was devised solely from investigation of the dialogues and is a qualitative technique. This technique is known as grounding theory (Glaser and Strauss, 1967) where categories are not imposed but derived from the data.

The second main part of the coding scheme related to talk which concerns the task. Watson and Sasse (1996) suggested that users will rate the quality of the technology differently depending on the task in which the participants are involved. For example, depending on whether the task is to learn a foreign language rather than simply report data, the users will perceive the quality of the technology differently. It is therefore important to understand the purpose of the sessions and to create categories to capture this information. It is essential to know how the system is being used. The task category is not unfamiliar in the literature (Carletta *et al.*, 1998; Bales 1950). Carletta *et al.* (1998) used an all-encompassing category to code task-relevant information, although they also utilised a separate category for any task-relevant utterances which included an information source. Extensive investigation of the TeachSpeech dialogues demonstrated that the current coding scheme should subdivide task into three components. The subdivisions of task include teaching talk, reporting talk and administration talk. Each of these different types of tasks were derived from the data and thus through an approximation of

grounding theory (Madill *et al.* 2000). These three categories will be described in turn.

Teaching talk is coded when tuition or instruction takes place as when the ESA receives support from the SLT concerning speech and language therapy techniques. As mentioned earlier, it has been reported that the type of task being completed and who completes the task, for example with non-native versus native speakers, affects whether video-mediated communication has benefits over audio-only communication (Veinott *et al.*, 1999; Reisberg *et al.*, 1987; Short *et al.*, 1976). It is thus important to determine the function of each session. Although all the interaction that takes place over the videoconferencing system has the potential to use both channels of information, i.e. audio and visual, interaction where teaching talk takes place may well elicit increased use of the visual channel. Consequently, it is useful to code this type of talk. The origins of this category were derived predominantly from the exploration of the TeachSpeech dialogues.

One of the main aims of the sessions between the SLT and ESA was to report to the SLT the progress of the child or children the ESA was supporting; thus a category labelled report talk was devised. This included all information regarding the child's well-being, particularly with respect to their speech or language impairment. The child may experience behavioural problems which are related to the speech and language problem and such feedback would also be included in this category. Analogous to the teaching talk classification, the origins of this category lie within the dialogues themselves.

The last category is termed administration talk. This included all talk which concerned managerial or administrative issues. Examples include arranging face-to-face visits, videoconferencing sessions and assessment appointments. This category is similar to the project management category in Olson and Olson's (1992) design meeting coding scheme. Their category was defined as: 'statements having to do with activity not directly related to the content of the design, in which people are assigned to perform certain activities, decide when to meet again, report on the activity (free of design content) from previous times and so on' (Olson and Olson, 1992, pg. 356). Here the assignment of tasks is generally a one way process. Generally, the SLT assigns tasks for the ESA to complete, although the SLT can also assign themselves tasks such as contacting other professionals. The main difference

between administration talk and project management is that reporting on an ongoing activity would not be included in the administration talk category. This category is closely related to Olson and Olson's (1992) project management category.

A preliminary analysis of four dialogues was performed to determine whether all the categories were necessary with no redundancy between any two categories. As a precautionary measure in the preliminary analysis, an 'other' category was included, where any turn which could not be allocated into one of the alternative categories was placed. However, this was removed as all turns within the dialogues could be coded using the remaining categories.

To summarise, the coding scheme devised here used Carletta *et al.*'s (1998) scheme as a general outline and used Grounding Theory to develop several other coding categories. In line with Carletta *et al.*'s (1998) coding scheme, technology talk, social talk and task talk were all coded. Two categories from this coding scheme were not included in the current coding scheme since they were inapplicable. These related to utterances that concerned an information source or were due to external distraction. Unique to the present coding scheme was the addition of the evaluation talk category and the apportionment of the task talk component. The different types of task included teaching, reporting and administration.

4.6 Method

The current coding scheme has been applied to transcribed audio tape recordings of videoconferences from the TeachSpeech model of service delivery, as outlined in Section 1.2.1. The videoconferencing usually followed a particular format, where a typical scenario consisted of the TeachSpeech SLT at one terminal in the I CAN nursery and an ESA at another terminal in various mainstream schools in the Wiltshire area. Exceptions included two ESAs using the system from the same school at the same time, and instances in which parents and teachers would use the system to make contact with the SLT. While there was no specific agenda for the meetings, the aim of the vast majority of the sessions was to update the SLT on the progress of a child or children the ESA was supporting. The ESA would report what the child had been doing during the previous week and would highlight any problem areas or areas of accelerated learning. A re-evaluation of previously set goals and, in some cases, discussion concerning administration would take place. The sessions would

typically last between fifteen and twenty-five minutes.

4.6.1 Participants

One TeachSpeech SLT and the four ESAs that she supported formed the participant group. All participants' permission was received in order to audio record the sessions.

4.6.2 Equipment

Videoconferencing system As outlined in Section 2.3.1, the videoconferencing system utilised was a BT VC6000 videoconferencing system, using a digital bandwidth of 384 kb/sec formed by six basic rate channels at 64 kb/sec (i.e. three ISDN2 lines). This is a high quality system comprised of a 48 kilobit audio channel, a 28 inch colour monitor and camera capable of pan, tilt and zoom, with remote comparison and picture in picture facility. The frame rate of the video in the TeachSpeech project was within in the 6.25 to 12.5 frames per second band.

Audio Recording Device All the audio recordings were made by the SLT at the I CAN nursery using a Sony WMD6C Pro Tape Recorder device. Recordings were taken from September 1997 until May 1998. Midway through the recording of the sessions, the recording device was installed directly into the videoconferencing unit and the audio quality of the recordings improved.

4.6.3 Transcripts

Twenty-three sessions were recorded, six of which were unusable due to poor audio quality and thus seventeen were transcribed. The transcription was conducted by the researcher and entailed carefully noting each word that was spoken, dividing speaker's utterances into turns and noting each instance of overlapping speech. Each tape was listened to at least twice. Each 20 to 30 minute videoconference meeting took approximately 3 hours to transcribe.

Words, turns and episodes of overlapping speech have been defined and utilised differently by different researchers. The research here has been modelled mainly on the work of Boyle *et al.*, (1994). Words were defined as an individual lexical item perceived as having acoustic boundaries. Turns were defined as the time

during which a speaker appears to 'hold the floor'. For a full discussion of turns and turn-taking behaviour see Sacks *et al.*, (1974). Only episodes of overlapping speech which extended beyond the original floor holder's speech were considered as new turns, although all speech is noted in the transcripts. This means that backchannels, a brief response by the listener signalling agreement attention or understanding (Boyle *et al.*, 1994), were not counted as turns.

An instance of overlapping speech was noted whenever two or more participants were perceived to be speaking simultaneously, as defined by Rutter and Stevenson, (1997). Episodes of overlapping speech were coded to show where each such episode began and ended.

4.6.4 Reliability of the Coding Scheme

As stated previously, the issue of reliability is of central importance (Bakeman and Gottman, 1997). Before the seventeen transcripts were coded, a second coder was trained on the coding scheme in order to ensure that inter-rater reliability could be achieved. Two coders analysed two of the dialogues, the results showed an inter-judge agreement of 82%. The Kappa coefficient statistic (Siegel and Castellan, 1988) was calculated and a kappa of 0.85 was obtained (N=439, κ=2). This result shows significantly more agreement occurring than would be expected by chance. It was felt that reliability decay would not be a concern since the dialogues were then coded over a period of five consecutive days.

4.7 Results

To reiterate, the two aims of evaluating the process of communication were to describe the dialogues in terms of their content and structure and to assess the value of the visual channel. It should be noted that there are four three-party dialogues included in the analysis, thus when analysing the SLT and ESA separately there are seventeen SLT units of dialogue and twenty-one ESA units of dialogue.

4.7.1 Descriptions of the Dialogues Utilising Content Coding and Structural Analysis

Table 4.3 shows the percentage of turns classified within each of the coding categories. The majority of speech that takes place across the videoconferencing link is

Table 4.3: Percentage of turns allocated to each category averaged over the seventeen sessions where SD represents the standard deviation

Coding Category	Mean percentage of turns	Minimum	Maximum	SD
Technology	2.1	0.0	11.6	3.3
Social	12.1	3.1	29.7	7.9
Evaluation	2.2	0.0	19.1	4.9
Teaching	6.3	0.0	27.9	7.9
Reporting	56.8	33.3	79.2	17.8
Administration	20.5	0.0	48.1	13.0

Table 4.4: Percentage of turns allocated to each category in each session. Here Eval represents evaluation talk and Admin represents Administration talk.

Session	Technology	Social	Eval	Teaching	Reporting	Admin
1	3	5	6	0	77	9
2	2	3	1	1	74	17
3	3	9	0	28	40	20
4	0	23	3	7	38	29
5	7	6	0	8	34	45
6	4	5	1	9	71	10
7	0	18	8	2	51	21
8	0	16	6	0	59	17
9	0	15	0	8	73	4
10	2	9	0	21	45	23
11	0	7	0	11	33	47
12	1	6	10	0	35	48
13	0	7	5	3	79	6
14	2	25	0	0	69	4
15	0	30	0	0	38	32
16	0	14	0	0	74	12
17	12	7	0	6	75	0

Table 4.5: Structural analysis of dialogues, averaged over the seventeen sessions where OS represents the term overlapping speech

Unit Analysed	Mean	Minimum	Maximum	SD
Words	1881.2	571.0	5298.0	1109.5
Turns	119.5	42	360.0	72.9
Words per turn	15.5	5.6	23.5	4.5
SLT Words per turn	13.8	8.2	21.2	3.6
ESA Words per turn	16.9	6.5	36.9	7.1
Number of times OS occurred	44.5	22.0	96.0	17.7

reporting talk (56.8%), with administration being the second largest component of the sessions (20.5%). A mean of 6.3% is reported for teaching talk. This shows that the link can be and is used for this task. The majority of non-task talk is social talk (12.1%), with technology and evaluation talk being virtually identical at 2.1% and 2.2%, respectively. The finding that only 2.1% of the turns are categorised as technology talk indicates that the technology is not intrusive and does not disrupt the sessions significantly.

From the results, shown in Table 4.4, it can also be seen that the dialogues are relatively similar. There is always reporting talk, social talk and administration talk, except for dialogue 7 where there is no administration talk. It is also noteworthy that the technology talk is consistently low across all dialogues. A general pattern does seem to emerge.

Comparison of words, turns and words per turn was previously modelled by Boyle *et al.*, (1994). Structural analysis is another way of capturing the communication process. The purpose here is to examine the length of the dialogues and to obtain an indication of the interactivity of the dialogues, as can be measured by words per turn and amount of overlapping speech. As was outlined previously in Section 1.4, structural analysis can provide informative data which may or may not confirm users' subjective comments.

Table 4.5 shows the turns, words, words per turn and the number of times overlapping speech occurred in the dialogues, where the results have been averaged across the seventeen dialogues. Clearly, the amount of words and turns will dif-

Table 4.6: Structural analysis of dialogues, illustrated for every dialogue. Where words per turn are represented by WPT and number of times overlapping speech occurred by TO.

Session	Words	Turns	WPT	SLT WPT	ESA WPT	TO
1	5298	360	14.72	8.22	21.57	96
2	571	43	13.28	16.76	8.41	26
3	2691	160	16.82	10.45	16.20	38
4	1175	116	10.13	13.34	6.53	31
5	2068	111	18.63	13.36	18.14	38
6	963	79	12.19	14.35	10.23	22
7	2843	170	16.72	21.17	13.92	68
8	1386	78	17.77	16.97	16.56	40
9	2710	145	18.69	16.81	13.04	54
10	1681	128	13.13	13.24	14.20	32
11	2031	144	14.10	10.54	16.58	50
12	986	42	23.48	14.08	26.28	39
13	1321	61	21.66	19.71	36.88	54
14	2001	101	19.81	13.73	20.12	54
15	831	75	11.08	10.66	12.25	33
16	1536	96	5.58	10.67	18.00	43
17	1888	122	15.47	10.30	18.80	38

fer depending on the task of the meeting. For example, Daly-Jones *et al.* (1998) transcribed dialogues from video recordings of two-party videoconferencing sessions where participants performed a collaborative problem-solving task. They recorded an average of 165.8 turns per dialogue. Results were not quoted for mean number of words per dialogue. Certainly, a direct comparison to this study would be spurious. However, words per turn are not influenced by the length of the task and indicate to some extent how interactive a session is. Daly-Jones *et al.* (1998) reported 13.9 words per turn for their high quality videoconferencing sessions. Table 4.6 shows that words per turn are 15.5, a similar finding. Another study which examined words per turn is O'Conaill *et al.* (1993) who recorded four-party meetings which took place over a high quality videoconferencing system, the primary purpose of which was for information exchange. They reported 19.23 words per turn for their high quality videoconferencing sessions (LIVE-NET) and 17.08 for their face-to-face condition. Daly-Jones *et al.* (1998) proposes that short, frequent turns and overlapping speech is indicative of verbal fluency. Therefore the relatively short mean turn length in this study suggests greater verbal fluency in relation to the O'Conaill

et al. (1993) study. However, comparisons with previous studies are difficult due to differences in the type of task and participant numbers. Table 4.6 shows the results for all seventeen dialogues.

The results here show that on average there are 44.5 instances of overlapping speech per dialogue. Boyle *et al.* (1994) analysed face-to-face dialogues using the map task and analysed the number of times overlapping speech occurred. They were investigating how dialogues of partners who could see each other differed from those who could not see each other. The number of times overlapping speech occurred when partners could see each other was 11.52 times per dialogue and this increased for those partners who could not see each other. Ideally, the percentage of times overlapping speech occurred would be quoted here, unfortunately percentages were not quoted in the Boyle *et al.* (1994) paper. However, an accurate indication of dialogue size is given by the average number of words and turns in a dialogues. The Boyle *et al.* (1994) study states that in dialogues where participants could see each other an average of 1049 words were spoken in an average of 142.5 turns. Thus, it would seem that these dialogues are longer than the TeachSpeech dialogues and that the amount of overlapping speech, if the dialogues were to have the same characteristics, should therefore be higher. However, this is not the case.

Boyle *et al.* (1994) suggested that low numbers of overlapping speech indicated smooth turn taking. However, simultaneous speech can either be taken to indicate a problem in floor control or it may be taken to indicate the degree of interactivity that is occurring within a dialogue (Daly-Jones *et al.*, 1998; Sellen, 1992; Rutter and Stephenson, 1977). Rutter and Stephenson (1977) suggest that high numbers of interruptions indicate comfort in a conversation. They found that face-to-face conversations have the most interruptions while audio-only conversations have the least and videoconferences fit somewhere in between. It is important to note that, as defined previously, the overlapping speech noted in the TeachSpeech dialogues includes interruptions. Sellen (1992) uses questionnaire results, speaker switching time analysis and an analysis of the structure of the dialogues in forming the conclusion that less simultaneous speech is indicative of a more formal type of interaction.

Comparison between studies is difficult since the definition of overlapping speech is unclear and many studies do not provide raw data and thus the percent-

Table 4.7: Illustrates which ESAs contributed to which time zones and how much experience they had. Experience is measured by the number of times the ESA had taken part in a videoconference prior to each time zone

ESA	Time 1		Time 2	
	Experience	Number of sessions	Experience	Number of sessions
ZN	1	4	18	4
MK	11	3	25	2

age of overlapping speech cannot be calculated. However, episodes of overlapping speech in this study seem high, potentially indicating a more interactive style of communication. This conclusion concurs with the findings from the words per turn analysis which also hints at an interactive form of communication taking place in the dialogues.

Do the Dialogues Change Over Time?

The second issue which was investigated was whether or not dialogue content and structure change over time. It was decided that the dialogues would be split into two terms with the first extending from July 1997 until December 1997 and the second from January 1998 until April 1998. Thirteen dialogues were included in this analysis as three did not fall into this time period. These time periods were chosen since both included the beginning and end of term; sessions which generally consisted of a great deal of administration talk. Table 4.7 clarifies which ESAs took part in which time period and establishes the amount of experience each ESA had prior to the start of each time period. There is only one SLT, thus every session took place with an ESA and the same SLT. At the start of Time 1, the SLT had taken part in 12 sessions and at the start of Time 2 she had taken part in 47 sessions. These figures do not include training time or any demonstrations the SLT may have given.

Table 4.7 Illustrates which ESAs contributed to which time zones and how much experience they had. Experience is measured by the number of times the ESA had taken part in a videoconference prior to each time zone.

Of interest was whether the dialogues would become more interactive over time. Words, turns and words per turn were examined. Table 4.8 shows the means

Table 4.8: Means of words, turns and words per turn over time where min is the minimum number of words and max is the maximum number of words. OS represents overlapping speech.

	Time 1				Time 2			
Unit Analysed	Mean	Min	Max	SD	Mean	Min	Max	SD
Words	2179.0	571.0	5298.0	1459.0	1611.0	831.0	2843.0	551.5
Turns	139.0	43.0	360.0	96.7	102.0	42.0	170.0	41.8
Words per turn	14.5	5.6	18.7	4.7	16.6	11.1	23.5	4.7
Number of times OS occurs	45.8	26.0	96.0	21.9	43.3	22.0	68.0	14.2

Table 4.9: Mean number of turns and words split over time for each category of interest

Coding	Time 1				Time 2			
Category	Mean	Min	Max	SD	Mean	Min	Max	SD
Reporting	774.7	213.0	1466.0	439.1	1089.9	615.0	2074.0	569.5
Social	84.6	8.0	363.0	127.9	130.7	36.0	350.0	110.1
Teaching	85.1	0.0	246.0	99.9	127.7	0.0	404.0	153.1
Technology	43.4	0.0	133.0	46.1	15.1	0.0	56.0	15.1

that were obtained.

Paired samples t-test were calculated between Time 1 and Time 2 for words, turns and turn length. No significant differences were found. The overall structure of the dialogues did not differ over time suggesting that participants adopted a style of interaction early on in their use of the technology which remained unchanged with experience. While no differences were found in the analysis of the surface structure of the dialogues over time, of specific interest was whether the distribution of types of talk changed over time. Table 4.9 shows the means obtained for the number of words over Time 1 and Time 2.

Previous studies on gesture showed that interaction which is mediated by video adapts over time, (Rudman *et al.*, 1997; Rudman and Dykstra-Erickson, 1994). Paired t-tests were conducted on number of words within each category and results showed no significant differences. It can be seen from the means presented in Table 4.9 that the distribution of communicative topics remains remarkably stable over time.

To summarise, findings here conclude that over time little adaptation occurs

Table 4.10: Example of the use of the video channel - 1

Speaker	Extract	Coding
SLT	< so he's he's under understanding the difference he's just still confusing them when he says them / isn't he	Teaching
ESA	yeah yeah yeah >	Teaching
SLT	yeah so I think it might be best if you do if you turn the activity around so he's the teacher so he has to tell you which bit to colour now colour you know you're the teacher you tell me who's hair to colour or who's shirt to colour	Teaching
ESA	ok	Teaching
SLT	< or who to give or who to give this toy to if you can find a boy and girl doll or picture or something they / can	Teaching
ESA	sure sure >	Teaching
SLT	< yeah but a dressing game is very helpful do you have any any games where you dress the person / shall I shall I show you one I'll show you one just a minute Jane this is a bench game can you see those	Teaching
ESA	mhm mhm yes yes yes >	
SLT	a boy a boy and a girl, it's like paper dolls in a way or and then there's a selection of clothing	Teaching
ESA	mhm	

and that both the structure and the communicative topics seem to remain constant. From this, it seems that users adapt quickly to the technology and may only require a small time period to adjust to communicating across the videoconferencing link.

4.7.2 The Use of the Visual Channel

The transcripts were examined for instances where a speaker explicitly utilises the visual channel. There were three occasions noted in the transcripts where the SLT or ESA described sounds to the other. Another use of the visual channel is illustrated on Table 4.10. The SLT shows the ESA how to conduct a speech and language exercise. This was recorded once in the transcripts. It was also noted from the transcripts that participants used the video channel to view pictures and work the children had completed. This occurred twice and Table 4.11 illustrates one of the examples. Thus, the visual channel was explicitly used 6 times in the 17 sessions recorded.

Table 4.11: Example of the use of the video channel - 2

Speaker	Extract	Coding
ESA	mhm What else have we done so far oh can you see the picture	Teaching
SLT	yes yes	Teaching
ESA	Michelle drew that	Teaching
SLT	oh wonderful	Teaching
	boy and girl doll or picture or something they / can	
ESA	and and Neil put the tree on	Teaching
SLT	fantastic	Teaching
ESA	and I asked him where the sky goes on the picture and he	
	got it right up at the top	
SLT	yeah	

4.8 Discussion

Monk *et al.* (1996) state that in order to understand and draw conclusions from an evaluation, a wide variety of measures should be obtained. The results from the communication analysis contribute a great deal to the overall understanding of the TeachSpeech evaluation. These objective measures provided detailed information concerning the use of the videoconferencing system. Description of the dialogues, in terms of content coding, served to answer three specific issues. The questions of interest include whether the dialogues contained an intrusive amount of interaction concerning the technology, if the technology was able to support an adequate amount of social interaction and if the quality of the present system was sufficient to sustain more complex interactions such as teaching. Finally, the following discussion explores further whether or not communication changes over time and whether or not explicit use of the visual channel occurred.

A great deal of discussion around the use of technology would suggest either a problem in usability or an intrusion of technology, which would be apparent in the communication between the SLT and the ESAs. As can be seen from the results, only an average of 2.1% of turns were used to talk about the technology with little variability between dialogues. This adds strength to the overall finding and supports the supposition that the technology is not complicated or intimidating to use. This result was highly encouraging, suggesting that participants are largely able to ignore the technology while conducting their meetings. Olson *et al.* (1997) categorised

technology confusion and technology management as part of their analysis. While they do not quote any results pertaining to these two categories, they do illustrate their results diagrammatically and from this it can be seen that technology confusion and management certainly do occur in the work of Olson *et al.* (1997). Whether or not the amount of technology talk in the present study exceeds that of the Olson study cannot be determined with the data they present.

It is clearly essential that technology talk does not interfere with the communication process, but it is also important that the technology actually sustains social interaction. Social interaction can be a vital component in a working relationship, especially one which requires close co-operation. Gutek (1997) believes that social talk is an important part of building a friendly and long term relationship. Technology therefore needs to sustain social communication (Boden and Molotch, 1994). As can be seen from the content analysis, 12.1% of turns are social in nature indicating that this technology does support social interaction. This is the largest component of the non-task talk and is approximately 6 times greater than both of the other categories within the non-task talk section. Thus, in relative terms, this may be interpreted as a merit of the videoconferencing system, since it supports social interaction while rarely interfering with topic of discussion. In the Carletta *et al.* (1998) study, dialogues were coded for social information. Their results showed that 10% of turns could be classified in the social category. The authors suggest that this is low but do not provide the benchmark against which they compared this figure. In the questionnaires which evaluated stakeholders' perceptions, the stakeholders are specifically asked whether they feel able to chat socially. Their subjective response suggested that this type of interaction was possible. The analysis conducted here confirms this and adds considerable strength to the overall finding.

The technology satisfies two of the requirements discussed so far. However, does it support complex interaction such as the SLT teaching the ESAs new speech and language techniques? From the results, it can be seen that on average 6.3% of the interaction was classified as teaching talk. The highest percentage of talk within any session was 29% indicating, on occasions, quite extensive amounts of teaching did take place. Therefore, teaching can be supported across the videoconferencing link and this allows the SLT to teach any relevant speech and language therapy techniques to the ESA who is untrained in this area. Other research findings indicate that

effective teaching may take place over videoconferencing links (Jacob and Rodgers, 1997). Jacobs and Rodgers (1997) used basic rate ISDN, which only uses two lines, to teach English to a group of students in Belgium. Their results indicated that effective learning had taken place. The present study uses ISDN with 6 lines thus obtaining an overall higher level of quality, but again we have no real benchmarks to compare against, and thus this warrants further study.

4.8.1 Do the Dialogues Change Over Time?

The findings here indicated that there seemed to be no effect of experience with the videoconferencing system. In general, there is little research on the effects of experience on video-mediated communication. Newlands *et al.* (in press) investigated how participants adapted to text-based computer mediated-communication (CMC). Participants completed different versions of the map task over three consecutive days and their dialogues were coded using CGA. This analysis showed that dialogues did indeed change over time as by the final task, the dialogues in the CMC condition contained a different pattern of moves than they had initially. They also found that the CMC dialogues became shorter over time. Newlands suggests that the adaptation to novel communication modes does occur and that it occurs swiftly. Rudman *et al.* (1997) believe that many studies which employ short-term laboratory experiments do not reveal the true nature of video-mediated communication. They believe that studies where the participants do not have a vested interest, or the users have only limited exposure to the system, will not yield valid results.

The TeachSpeech sessions that were transcribed spanned a nine-month period and those involved certainly had an interest in the outcome. The analysis explored the constancy of the content and structure over time and the findings showed that the sessions remained remarkably stable over time. Topics tackled remained similar, as did the length and interactivity of the dialogues. The two aspects of the dialogue investigated here show no change over time or experience. From this, it is found that users seem to adapt quickly to the system. Thus, it could be suggested that new users are likely to find the system easy to use and, as a consequence, more people are likely to utilise the technology. However, it should be noted that there were only 7 dialogues in Time 1 and 6 in Time 2. Possibly a larger sample size may be required to draw firm conclusions.

4.8.2 Explicit use of the visual channel

The extent to which the visual channel was explicitly used was also investigated. Several explicit uses of the visual channel were found. These included the mimicking of sounds children produce which utilises lip-movement to aid comprehension, illustrating how to conduct a speech and language exercise and to view pictures and work the children had completed. The videoconferencing tool was also used to observe children in their environment. For example, the ESA could virtually meet and observe a child before they started attending their school or the SLT might observe the child in the classroom setting. Unfortunately, none of these sessions was available for recording, but it is clear that the visual channel is very important in these instances. In these select instances, if the visual channel had not been present it is possible that the ESA might have had to visit the school in order to meet the child and this would have resulted in time lost in the classroom in addition to the travel costs.

Clearly, it would have been desirable to have comparable telephone meetings and face-to-face meetings for comparison. This would have allowed the direct comparison of results obtained from content coding and structural analysis. However, due to the nature of the study this was unfeasible as discussed in Section 1.1.3. Since these dialogues were not available, no attempt to suggest whether differences would have occurred in the results of the content or structure of the dialogues is made here. However, the task of the evaluation is to assess whether or not the proposed technology utilised here can be effective in providing support to ESAs and not to assess the relative merits of other systems and in that sense the results appear promising.

4.9 Conclusions

As can be seen, the content analysis of communication from the videoconference sessions is a useful research and evaluation tool. It has provided critical and reliable information about the use of the videoconference sessions. In the previous chapter, results obtained from questionnaires and interviews suggested that users were satisfied with the videoconferencing sessions and that they felt the video and audio channels were necessary and provided a useful way to communicate. The communication analysis seems to have confirmed this in a number of ways. Firstly, there is

infrequent talk concerning the technology, indicating that technological problems are rare and suggesting that users were able to concentrate on the aim of the sessions. Secondly, it was found that social interaction was supported by the technology, which is important in building good working relationships. Thirdly, teaching occurred across the link and other research (Jacobs and Rodgers, 1997) has shown that teaching is certainly feasible using this delivery mechanism. Another issue that was examined included participants' experience. The investigation of experience showed little effect of time on the structure or content of the dialogues, and it seems likely that users adapt quickly when using the videoconferencing technology. Finally, various examples attained from the transcription explicitly illustrated the need for a visual channel. The cases observed were infrequent yet they seemed important in the sense that, without the videoconferencing technology, the intended message could not have been conveyed effectively. The evaluation of the process of the communication that occurs between the SLT and ESA over the video link has added to the variety of techniques used to evaluate this service delivery model and from this analysis alone a detailed picture has emerged.

Chapter 5

Conclusions

5.1 Conclusions

The purpose of this thesis was to carry out an extensive evaluation of the telemedical project TeachSpeech. The TeachSpeech Project was a collaborative project between BT and I CAN, with support from the DTI. This innovative pilot project was established to ascertain the effectiveness of a service delivery model of therapy which utilised a videoconferencing system. The system was used by speech and language experts to support schools which have children with speech and language impairments. The goal of the evaluation was to assess the impact of the TeachSpeech project compared to the traditional model of therapy on measures of performance, perception and process.

A review of service delivery models within the speech and language therapy domain and other telemedical projects demonstrated that criticisms mainly concerned the lack of wide-ranging evaluation studies being conducted. Evaluations that were conducted tended to primarily focus on clinical effectiveness measures. Literature from the field of video-mediated communication demonstrated that evaluating the process of video-mediated dialogues was a useful and effective research tool. Keeping these findings in mind, the evaluation proposed and presented in Chapters 2, 3 and 4 involved a multifaceted approach which aimed to provide a range of measures of the impact of the TeachSpeech project. To select one evaluation measure and apply it to a specified topic and then to draw conclusions from the data produced seemed unsatisfactory. In contrast in the present study many measures were used the conclusions drawn can be considered with more confidence.

The TeachSpeech evaluation uses a variety of techniques to assess the performance, the perceptions of stakeholders and the process of communication. Techniques used to assess performance included a clinical tool termed the Enderby Outcome Measure and log sheets, which were used to keep a detailed account of how the therapists spent their time. The costs of each service delivery model were assessed by collecting information from various sources. Stakeholders' perceptions were assessed by canvassing head teachers, teachers, parents, ESAs and SLTs who completed questionnaires and gave feedback in semi-structured interviews. Lastly, a new coding scheme to assess the communication process in technologically mediated meetings was developed and was presented in Chapter 4. This coding scheme was applied to the audio recordings of the video conferencing sessions between the TeachSpeech SLT and the ESAs. The coding scheme was utilised to address issues concerning which communicative topics were encountered over the video link, to establish whether or not there was a need for the visual channel and the effects of experience.

The main conclusions concerning the performance of the two models of therapy follow. The clinical outcome measures were pivotal to the evaluation of the TeachSpeech project. They demonstrated that both groups of children were at a similar level of impairment at the beginning of the project, both groups significantly improved over time and both were at a comparable level at the final assessment. The differing modes of service delivery, therefore, achieved equally successful clinical outcomes.

A limitation of the design of the TeachSpeech project is the lack of a no treatment condition plus the restricted sample size. The log sheets indicated that there was no difference in the amount of time the therapist spent on a child although differences were found within the different components noted on the log sheets. It could be hypothesised that the amount of time the therapists spent on each child was constant, however from examination of the wide-ranging standard deviations obtained and interview data, it seems that therapists allocate their time according to the severity of a child's' impairment. In the traditional model of therapy the log sheets indicated more travel time, preparation time and face-to-face contact with the children as compared to the TeachSpeech model of therapy. The TeachSpeech group spent more time liasing than the comparison group and, although there was

no overall time saving, the way in which the therapists spent their time is important since the therapist can use their time in a more or less productive manner.

It was demonstrated that the TeachSpeech model of therapy was more expensive than the comparison model of therapy. It is pertinent to consider several factors: firstly, prices of videoconferencing technology are continually falling and secondly, ESAs were being trained in the TeachSpeech model of therapy and this cost was not included in the analysis because of assessment difficulties. With regard to all three performance criteria listed above, it should be noted that additional measures could have been taken to further investigate performance of the models of therapy. For example, it would have been desirable to use a battery of tests to assess the children's impairment and a measure of how much time the children spent with the ESA. Mainly due to time and financial constraints, these measures were not recorded.

The main findings from the analysis of the perceptions of stakeholders are as follows. The assessment of stakeholders' views revealed positive attitudes towards the TeachSpeech model of therapy. Parents were satisfied with the guidance their child was receiving and ESAs felt more confident in delivering support to children in the TeachSpeech model of therapy. Overall, users were satisfied with the usability of the videoconferencing system. These findings from the SLTs and ESAs were strengthened by results obtained from the coding scheme used to explore videoconferencing. The small amount of talk concerned with the technology indicated few technological problems. It was also found that social interaction took place which has been suggested as an important part of building good working relationships.

The role of the technology and its evaluation follow. The findings from the content coding showed that complex interaction, such as teaching, took place. There were several explicit observations that suggested there was a benefit from a visual channel, such as using the video link to demonstrate techniques and observe children. The effects of experience of using the videoconferencing system over an extended time period of several months was examined. The results showed that there was little difference over time in the structure or communicative topic of the dialogues. The analysis showed that the dialogues remained remarkably stable over time. A basic limitation of the design of the project rendered the recording of equivalent face-to-face dialogues difficult and since audio-conferencing was not utilised, no

comparison could be made with either face-to-face or audio-only dialogues. The discussion concerning the process of communication is therefore limited to comparison between results obtained here and benchmarks obtained from the literature.

The overall findings of the evaluation are encouraging. The project was a small scale evaluation study with some necessary design limitations. The analysis however examines a number of different evaluation criteria which provided complementary evidence for the usefulness of this form of service delivery. These results have implications for those interested in telemedical projects and those interested in video-mediated interaction. The evaluation of this innovative pilot project demonstrates the potential feasibility of this method of service delivery for children with speech and language impairments and it also highlights practical issues.

5.2 Further Possible Exploration of the TeachSpeech Evaluation

Due to the promising nature of the results, there are various areas of the TeachSpeech project that could be explored without changing the overall design of the project. Since the TeachSpeech project was a pilot study, a replication study would be interesting in which a larger sample size is used.

As the TeachSpeech project was designed by SLTs who were concerned with optimal care levels, only video-mediated communication was explored. However, it would also be useful to compare video-mediated dialogues that occurred between the TeachSpeech SLT and the ESAs with audio recording of audio-only sessions. or as mentioned previously face-to-face sessions. The analysis of audio-only sessions would demonstrate whether or not the dialogues of the users differed when the visual channel was removed and more generally how users would respond to such a system. Either attitudes or more importantly effectiveness might be compromised. Analysis of face-to-face meetings could provide insight into the possible effects, positive or negative, of utilising a videoconferencing system to communicate. Lastly, it would be revealing to examine and analyse audio-recorded sessions of the ESA interacting with the children in each model of therapy. This would indicate whether the type of support the ESAs were provided with translated into any differences in the interaction between the child and the ESA. Here the design of the project would need to

be modified as the ESAs utilised were the same for both groups of children.

5.3 Other Avenues

Since the overall findings of the evaluation were encouraging, there are numerous paths that future work could explore. For example, the video image could be used to supply images to aid the communication process or specific training could be given to the ESAs over the videoconferencing link. Finally, trials could be run to assess how effective video-mediated communication between the child and the therapist would be. Each of these possibilities will be considered in turn.

Whittaker and O'Conaill (1997) distinguish between the use of video to display the remote participants and the use of the video to display visual information. The latter is referred to as video-as-data. Anderson *et al.* (2000) assess the relative merits of video-as-data compared to the video image of a remote participant. Their results showed that participants rated video-as-data as more useful than the video image of the remote participant. Possibilities in this area exist for the TeachSpeech project. Images such as an internal representation of the mouth and throat may be useful in order to explain aetiologies of impairments to the ESA, such as cleft lip or cleft palate. An example of the use of video-as-data within the speech and language therapy domain is CLEFTNET, a telemedical project which enables SLTs to access electropalatography (EPG) (Gibbon *et al.*, 1998). EPG has proven effective in the treatment of articulatory disorders associated with cleft palate (Gibbon and Hardcastle, 1989). SLTs acquire EPG and acoustic data at remote sites and send the information electronically for interpretation. Video-as-data therefore may be a helpful additional tool to explain and describe how to correct specific impairments.

Another possible use of the video technology within the scope of the TeachSpeech project is to train specifically the ESAs. Kingsnorth *et al.* (2000) evaluated the effectiveness of remote learning across videoconferencing systems. Videoconferencing equipment was installed in hospitals to deliver course modules, seminars and discussion groups to trainee doctors. Kingsnorth *et al.* (2000) found that the learning experience was effective and popular with doctors. The above study indicates that training across a videoconferencing link is a feasible and effective option. A practical addition of the uses of the videoconferencing system in the TeachSpeech project might be to train the ESAs or parents in therapeutic techniques.

An innovative use of the TeachSpeech videoconferencing system might be to encourage the use of the system for direct interaction between the therapist and the child. There are a few examples in the literature which illustrate the ability of children to interact directly over videoconferencing units. Fels *et al.* (1999) developed a system called PEBBLES (Providing Education By Bringing Learning Environment to Students). PEBBLES is a videoconferencing robot which allows school children to access their regular classroom while confined to hospital. Fels and Weiss (in press) surveyed sixty 5 to 12 year-olds and found that the children adapted with great ease to PEBBLES and were able to establish a strong rapport with the video unit and their absent classmates. Fels *et al.* (1999) found that the remote participant could communicate relatively successfully over the link. Dossetor (1999) examined the use of videoconferencing as an outreach service for children with psychiatric disorders. In one case where a child suffered from an eating disorder, a SLT interacted directly with the child to advise on swallowing exercises. Unfortunately, no information was provided regarding the child's age or the outcome of the session. Overall users' perceptions indicated very high satisfaction levels and they felt that sensitive interviewing was possible through videoconferencing. Clearly, it is feasible for children to interact directly with a videoconferencing unit, as has been outlined in the above studies. Regarding exploration of the TeachSpeech project in this area, initial interactions between the therapist and child might involve informal greetings which could progress to short conversations. Once ease of communication was established over the link, further uses could be established.

TeachSpeech was a pilot project that was conducted in a naturalistic setting. The overall findings from the TeachSpeech evaluation are encouraging and suggest that the video link is useful within this speech and language therapy setting. In addition, the evaluation adds to other successful telemedical projects, perhaps touching on the future delivery of a whole range of different services. Wilson *et al.* (2000) envisages a hospital without walls. Perhaps in the future a similar concept might be in place providing children with speciality services within schools.

References

Allen, A., & Perednia, D. (1996). Telemedicine and the health care executive. *Telemedicine Today, 5*, 4–9.

Anderson, A. H., Mullin, J., Katsavras, E., McEwan, R., Grattan, E., & Brundell, P. (1999). Understanding multiparty multimedia interactions. In S. Brennan, A. Gibion, & T. D. (Eds.), *Proceedings of the American association for artificial intelligence, fall symposium, psychological models of communication in collaborative systems* (pp. 9–16). California: Menlo Park.

Anderson, A. H., O'Malley, C., Doherty-Sneddon, G., Langton, C., Newlands, A., Mullin, J., Fleming, A. M., & Velden, J. Van der. (1997). The impact of vmc on collaborative problem solving: An analysis of task performance, communicative process and user satisfaction. In K. E. Fin, A. Sellen, & S. B. Wilbur (Eds.), *Video mediated communication.* Hillsdale: Lawrence Erlbaum Associates.

Anderson, A. H., Smallwood, L., MacDonald, R., Mullin, J., & Fleming, A. M. (2000). Video data and video links in mediated communication: What do users value? *International Journal of Human-Computer Studies, 52*, 165–187.

Bain, B. A., & Dollagham, C. A. (1991). Clinical forum: Treatment efficacy, the notion of clinically significant change. *Language, Speech, and Hearing Services in Schools, 22*, 264–270.

Bakeman, R., & Gottman, J. M. (1997). *Observing interaction: An introduction to sequential analysis* (2nd ed.). Cambridge, UK: Cambridge University Press.

Bales, R. F. (1950). *Interactional process analysis: A method for the study of small groups.* Cambridge, MA: Addison-Wesley.

Bales, R. F. (1955). How people interact in conferences. *Scientific American*, 3–7.

Bales, R. F., & Borgatta, E. F. (1955). Size of group as a factor in the interaction profile. In A. P. Hare, E. F. Borgatta, & R. F. Bales (Eds.), *Small groups: Studies in social interaction.* New York: Alfred A Knopf.

Benschoter, R. A., Garetz, C., & Smith, P. (1967). The use of closed circuit tv and videotape in the training of social group workers. *Social Work Education Reporter, 15*, 18–20.

Bishop, D. V., & Edmundson, A. (1987). Language-impaired 4-year-olds : Distinguishing transient from persistent impairment. *Journal of Speech and Hearing Disorders, 52*, 156–173.

Boden, D., & Molotch, H. (1994). The compulsion of proximity. In R. Friedland & D. Boden (Eds.), *Nowhere: Space, time and modernity.* Berkeley, CA: University of California Press.

Boyle, E., Anderson, A. H., & Newlands, A. (1994). The effects of visibility on dialogue performance in a co-operative problem solving task. *Language and Speech, 37*, 1–20.

Broen, P. A., & Westman, M. J. (1990). Project parent : A pre-school speech program implemented through parents. *Journal of Speech and Hearing Disorders, 55*, 495–502.

Callahan, E., Hilty, D. M., & Nesbitt, T. S. (1998). Patient satisfaction with telemedicine consultation in primary care: Comparison of ratings of medical and mental health applications. *Telemedicine Journal, 4*, 363–369.

Carletta, J., Isard, A., Isard, S., Kowtko, J., Doherty-Sneddon, G., & Anderson, A. H. (1996). HCRC dialogue structure coding manual. *Human Communication Research Centre Publication.*

Carletta, J., Isard, A., Isard, S., Kowtko, J., Doherty-Sneddon, G., & Anderson, A. H. (1997). The reliability of a dialogue structure coding scheme. *Computational Linguistics, 23*, 13–32.

Carletta, J., McEwan, R., & Anderson, A. H. (1998). Communication in virtual supply chain teams. In *Proceedings of the tenth international ifip wg5.2/5.3 international conference prolamat 98.* Italy.

Chapanis, A. (1988). Interactive human communication. In I. Greif (Ed.), *Computer supported co-operative work: A book of readings.* Morgan Kamfrian.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In J. Levine, L. B. Resnik, & S. D. Behrend (Eds.), *Shared cognition: Thinking as social practice.* Washington, DC: APA Books.

Clarke, P. (1997). *I CAN teletherapy solution - specifying videoconferencing equipment for speech and language therapists.* Unpublished, British Telecommunications plc document.

Cohen, K. M. (1982). Speaker interaction: Video teleconferencing versus face-to-face meetings. In *Proceedings of teleconferencing and electronic communications* (pp. 189–199). Madison: University of Wisconsin.

Conrath, B. W., Dunn, E. V., & Higgins, C. A. (1983). *Evaluating telecommunications technology in medicine.* Dedham, MA: Artech House, Inc.

Cooper, J., Moodley, M., & Reynell, J. (1978). *Helping language development.* Edward Arnold.

Cooper, J., Moodley, M., & Reynell, J. (1979). The developmental language programme: results from a five year study. *British Journal of Disorders of Communication, 14,* 57–69.

Cukier, J. (1997). Cost-benefit analysis of telelearning: Developing a methodology framework. *Distance Education, 18,* 137 –152.

Cukor, P., O'Laughlen, J., & Bauer, L. (1995). Gazing in the crystal ball: Trends in video conferencing research and design. *Telemedicine Today, 3,* 9–18.

Daly-Jones, O., Monk, A., & Watts, L. (1998). Some advantages of video-conferencing over high-quality audio conferencing: Fluency and awareness of attentional focus. *International Journal of Human-Computer Studies, 49,* 21–58.

Delapain, C. B., Lindborg, C. E., Norton, S. A., & Hastings, J. E. (1993). Tripler pioneers telemedicine across the pacific. *Hawaii Medical Journal, 52*, 338– 339.

Doherty-Sneddon, G., Anderson, A. H., Garrod, S., O'Malley, C., Langton, S., & Bruce, V. (1997). Face-to-face and video mediated communication: A comparison of dialogue structure and task performance. *Journal of Experimental Psychology: Applied, 3*, 105 – 125.

Dossetor, D. R., Nunn, K. P., Fairley, M., & Eggleton, D. (1999). A child and adolescent psychiatric outreach service for rural new south wales: A telemedicine pilot study. *Journal of Paediatric Child Health, 35*, 525–529.

Dunn, E. V., Conrath, D. W., Bloor, W. G., & Tranquada, B. (1977). An evaluation of four telemedicine systems for primary health care. *Health Services Research, 12*, 19–29.

Dunn, L. M. (1981). *Peabody picture vocabulary test.* Circle Pines, MN : American Guidance Service.

Eiserman, W. D., McCoun, M., & Escobar, C. M. (1990). A cost-effectiveness analysis of two alternative program models for serving speech-disordered preschoolers. *Journal of Early Intervention, 14*, 297– 317.

Enderby, P. M. (1997). *Therapy outcome measures, user's manual.* Singular Publishing Group, Inc.

Fajardo, L. L., Yoshino, M. T., Seeley, G. W., Hunt, R., Hunter, T. B., Friedman, R., Cardenas, D., & Boyle, R. (1990). Detection of breast abnormalities on teleradiology transmitted mammograms. *Investigating Radiology, 25*, 1111–1115.

Fels, D. I., & Weiss, T. (In press). Towards determining an attention getting device for improving interaction during video-mediated communication. *Computers in Human Behaviour.*

Fels, D. I., Williams, L., Smith, G., Treviranus, J., & Eagleson, R. (1999). Developing a video-mediated communication system for hospitalised children. *Telemedicine Journal, 5*, 193– 208.

Fey, M. E., Cleave, P. L., Long, S. H., & Hughes, D. L. (1993). Two approaches to the facilitation of grammar in children with language impairment: An experimental evaluation. *Journal of Speech and Hearing Disorders, 36*, 141–157.

Field, M. J. (1996). *Telemedicine: A guide to assessing telecommunications in health care.* Washington, D. C.: National Academy Press.

Forsberg, D. A. (1995). Quality assurance in teleradiology. *Telemedicine Journal, 1*, 107 –114.

Franken, E. A. (1996). Teleradiology: Moving into the mainstream. *Telemedicine Today, 4*, 25–28.

Garshnek, V., Logan, J. S., & Hassell, L. H. (1997). The telemedicine frontier: Going the extra mile. *Space Policy, 13*, 37–46.

Gibbard, D. (1994). Parental-based intervention with pre-school language delayed children. *European Journal of Disorders of Communication, 29*, 131 – 150.

Gibbon, F., Crampin, L., Hardcastle, B., Nairn, M., Razzell, R., Harvey, L., & Reynolds, B. (1998). Cleftnet (Scotland): A network for the treatment of cleft palate speech using EPG. *International Journal of Language and Communication Disorders, 33*, 44 –49.

Gibbon, F., & Hardcastle, W. J. (1989). Deviant articulation in a cleft palate child following late repair of the hard plate: A description and remediation procedure using electropalatography. *Clinical Linguistics and Phonetics, 3*, 93–110.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research.* New York: Aldine.

Grisby, J. and Kaehny, M. M. . (1993). *Analysis of expansion of access care through use of telemedicine and mobile health services* (Report 1: Literature review and analytic framework No. 500-92-0046). Denver, CO: Department of Health and Human Services.

Gutek, B. A. (1997). Dyadic interaction in organisations. In C. L. Cooper & S. E. Jackson (Eds.), *Creating tomorrow's organizations: A handbook for future. research in organizational behavior.* Chichester: Wiley.

Heinicke, C., & Bales, R. F. (1953). Developmental trends in the structure of small groups. *Sociometry, 16*, 7–38.

Helge, D. (1983). *Images and trends in rural special education.* Washington University, National Rural Development Institute: Bellinham.

Hoffert, S. (1997). Telemedicine boom awaits cost studies. *Scientist, 11*, 24– 29.

Houghton, G., & Isard, S. (1987). Why to speak. what to say, and how to say it. In P. Morris (Ed.), *Modelling cognition* (pp. 249–267). London: Wiley.

House, A. M. (1993). Distance health systems Canadian collaboration. In *Proceedings of the Mayo telemedicine symposium* (pp. 46–55). Rochester: Mayo Clinic.

Institute of Medicine Committee. (1997). Telemedicine needs better evaluation. *Public Health Report, 112*, 5–6.

Jacobs, G., & Rodgers, C. (1997). Remote teaching with digital video: A transnational experience. *British Journal of Education Technology, 28*, 292–304.

Johansson, I., Magnusson, M., & Wallin, E. (1997). Videotelephony as a training tool for people with aphasia and mental retardation. In H. Murphy (Ed.), *Proceedings from technology for disabled people* (pp. 16–22). Los Angeles.

Justice, J. W., & Decker, P. G. (1979). Telemedicine in a rural health delivery system. *Advances in Biomedical Engineering, 7*, 101–171.

Katsavras, E. (2000). *Evaluation of the TeachSpeech project.* Unpublished report.

Khan, L. M. L., & Lewis, N. P. (1986). *Khan-lewis phonological analysis.* MN: Circle Pines.

Kingsnorth, A., Vranch, A., & Campbell, J. (2000). Training for surgeons using digital satellite television and videoconferencing. *Journal of Telemedicine and Telecare, 6*, 29– 31.

Knowles, W., & Masidlover, M. (1979). *Derbyshire language scheme.* Derbyshire Education Authority.

Kowtko, J. C. and Isard, S. and Doherty-Sneddon, G. (1991). *Conversational games in dialogues* (Tech. Rep. No. HCRC/RP-26 Publications). University of Edinburgh.

Kysela, G., Hillyard, A., McDonald, L., & Ahlsten-Taylor, J. (1981). Early intervention: Design and evaluation. In R. L. Schiefelbush & D. D. Bricker (Eds.), *Early language: Acquisition and intervention.* Baltimore: University Park Press.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159–174.

Law, J. (1997). Evaluating intervention for language impaired children: A review of the literature. *European Journal of Disorder of Communication, 32*, 1– 14.

Levin, H. M. (1983). *Cost-effectiveness a primer: new perspectives in evaluation* (vol. 4 ed.). Newbury Park, California: Sage.

Long, S. H., & Fey, M. E. (1989). *Computerized profiling (Version 6.2).* New York: Ithaca.

MacDonald, J. D., Blott, J. P., Gordon, K., Spiegel, B., & Hartmann, M. (1974). An experimental parent-assisted treatment program for pre-school language-delayed children. *Journal of Speech and Hearing Disorders, 39*, 395–415.

Madill, A., Jordan, A., & Shirley, C. (2000). Objectivity and reliability in qualitative analysis: Realist, contextualist and radical constructionist epistemologies. *British Journal of Psychology, 91*, 1–20.

McConkey, R. (1985). *Working with parents.* Cambridge, Massachusetts: Brookline Books.

McDade, H., & Varnedoe, D. (1987). Training parents to be language facilitators. *Topics in Language Disorders, 7*, 19 – 30.

McLaren, P., & Ball, C. J. (1995). Telemedicine: Lessons remain unheeded. *British Medical Journal, 310*, 1390–1396.

Monk, A., McCarthy, J., Watts, L., & Daly-Jones, O. (1996). Measures of process. In Requirements & Evaluation (Eds.), *Thomas, C.* Springer.

Moore, G. T., Willemain, T. R., Bonanno, A. B., Clark, W. D., Martin, A. R., & Mogielnicki, R. P. (1975). Comparison of television and telephone for remote medical consultation. *The New England Journal of Medicine, 13*, 729–732.

Morley, I. E., & Stephenson, G. M. (1997). *The social psychology of bargaining.* London: George Allen Unwin.

Muller, C., Marshall, C. L., Krasner, M., Cunningham, N., Walllerstein, E., & Thomstad, B. (1977). Cost factors in urban telemedicine. *Medical Care, 15*, 251 –259.

Mun, S. K., Horri, S., & Benson, H. (1989). Experience with image management networks at three universities: Is the cup half-empty of half-full? In *Medial imaging III: Pacs systems design and evaluation. Proceedings of the society of photo-optical instrumentation engineers* (pp. 194–201).

Newlands, A. (1998). The effects of computer mediated communication on the process of communication and collaboration. *Unpublished PhD Thesis, Glasgow University, Department of Psychology.*

Newlands, A., Anderson, A. H., & Mullin, J. (In press). Adapting communicative strategies to computer-mediated communication: An analysis of task performance and dialogue structure. *Journal of Experimental Psychology.*

Nickelson, D. (1998). Telehealth and the evolving healthcare system: Strategic opportunities for professional psychology. *Professional Psychology: Research and Practice, 29*, 1000–1000.

O'Conaill, B., Whittaker, S., & Wilbur, S. (1993). Conversations over video conferences: An evaluation of the spoken aspects of video-mediated communication. *Human-Computer Interaction, 8*, 389 –428.

Olson, G. M., & Olson, J. S. (1992). Defining a metaphor for group work. *IEEE Software, 9*, 93– 95.

Olson, G. M., Olson, J. S., Carter, M., & Storrosten, M. (1992). Small group design meetings: Analysis of collaboration. *Human-Computer Interaction, 7*, 347 – 374.

Olson, J. S., Olson, G. M., & Meader, D. K. (1997). Face to face group work compared to remote group work with and without video. In K. E. Fin, A. Sellen, & S. B. Wilbur (Eds.), *Video mediated communication.* Hillsdale: Lawrence Erlbaum Associates.

Perednia, D. A., & Allen, A. (1995). Telemedicine technology and clinical applications. *Journal of the American Medical Associations, 273,* 483–488.

Poole, M. S., & Hirokawa, R. Y. (1986). *Communication and group decision making.* New York: Sage.

Power, R. J. D. (1979). The organisation of purposeful dialogues. *Linguistics, 17,* 107 - 152.

Preston, J., Brown, P. W., & Hartley, B. (1992). Using telemedicine to improve health care in distant areas. *Hospital and Community Psychiatry, 43,* 25– 31.

Reisberg, D., McLean, J., & Goldfield, A. (1987). A lip-reading advantage with intact auditory stimuli. In B. Dodd & R. Campbell (Eds.), *Hearing by eye: The psychology of lip-reading.* London: Lawrence Earlbaum Associates.

Renfrew, C. E. (1986). *Action picture test.* Oxford: University of Oxford Press.

Reynell, J. K. (1983). *Reynell developmental language scales.* Windsor: NFER-Nelson.

Roulstone, S., Glogowska, P., Enderby, P., & Peters, T. J. (1999). Issues to consider in the evaluation of speech and language therapy for pre-school children. *Child: Care, Health and Development, 25,* 141 - 155.

Rudman, C., & Dykstra-Erickson, E. (1994). Adaptation over time is required for effective use of video for group work. In *CSCW'94 : Workshop on video-mediated communication.*

Rudman, C., Hertz, R., Marshal, C., & Dykstra-Erickson, E. (1997). Channel overload as a driver for adoption of desktop video for distributed group work. In K. E. Finn, A. Sellen, & S. B. Wilbur (Eds.), *Video mediated communication.* Hillsdale: Lawrence Earlbaum Associates.

Ruscello, D. M., Cartwright, L. R., Haine, K. B., & Shuster, L. I. (1993). The use of different service delivery models for children with phonological disorders. *Journal of Communication Disorders, 26*, 193–203.

Rutter, D. R. (1987). *Communicating by telephone.* Oxford: Pergamon Press.

Rutter, D. R., & Stephenson, G. M. (1977). The role of visual communication in synchronising communication. *European Journal of Social Psychology, 7*, 29–37.

Sacks, H., Schegloff, E. A., & Jefferson, G. A. (1974). A simple schema for the organisation of turn-taking in conversation. *Language, 50*, 696– 735.

Schegloff, E. A., Jefferson, G. A., & Sacks, H. (1977). The performance for self-correction in the organisation of repair in conversation. *Language, 53*, 361–382.

Scott, W. W., Rosenbaum, E. J., Ackerman, S. J., Reichle, R. L., Magid, D., Weller, J. C., & Gitlin, J. N. (1993). Subtle orthopaedic fractures: Teleradiology workstation versus film interpretation. *Radiology, 187*, 811– 815.

Sellen, A. J. (1992). Speech patterns in video-mediated conversations. In P. Bauersfeld, J. Bennett, & G. Lynch (Eds.), *Proceedings of computer and human interaction '92* (pp. 49–59). New York: ACM Press.

Sellen, A. J. (1995). Remote conversations: The effects of mediating talk with technology. *Human-Computer Interaction, 10*, 401 –444.

Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications.* London: Wiley.

Siegel, S., & Castellan, N. J. (1988). *Nonparametric statistics for the behavioral sciences* (2nd ed.). McGraw-Hill.

Slovis, T. L., & Guzzardo-Dobson, P. R. (1991). The clinical usefulness of teleradiology of neonates: Expanded services without expanded staff. *Pediatric Radiology, 21*, 333–335.

Snyder-McLean, L., & McLean, J. E. (1987). Effectiveness of early intervention for children with language and communication disorders. In M. Garlnick &

F. C. Bennett (Eds.), *The effectiveness of early intervention for at risk and handicapped children* (pp. 213–274). Academic Press, Inc.

Stamm, B. H., & Perednia, D. A. (1998). *Evaluating telehealth systems: Questions, frameworks, and templates.* Unpublished manuscript.

Tang, J. C., & Isaacs, E. (1993). Why do users like video? *Computer Supported Co-operative Work, 1*, 163– 196.

Taplin, P. S., & Reid, J. B. (1973). Effects of instructional set and experimenter influence on observer reliability. *Child Development, 44*, 547–554.

Tichenor, J., Balch, D., & Gustke, S. (1996). Operational issues at the east carolina university school of medicine. *Telemedicine Today, 4*, 28–29.

U.S. Department of Commerce. (1997). Telemedicine report to congress. *Available: http://www.ntia.doc.gov, 14.12.00.*

Veinott, E. S., Olson, J. M., Olson, G. M., & Fu, X. (1999). Video helps remote work: Speakers who need to negotiate common ground benefit from seeing each other. In *Proceedings of computer and human interaction '99.* Pittsburgh, PA: ACM Press.

Wade, B. (1994). Telemedicine. *National Conference of State Legislatures, 2*, 41–48.

Walters, T. J. (1996). *Deployment telemedicine: The walter reed army medical center experience.* Military Medicine.

Watson, A., & Sasse, M. A. (1996). Evaluating audio and video quality in low-cost multimedia conferencing systems. *Interacting with Computers, 8*, 225–275.

Waxler, N. E., & Mishler, E. G. (1966). Scoring and reliability problems in interaction process analysis. *Sociometry, 29*, 28– 40.

Weinstein, R. S., Bloom, K. J., & Rozek, L. S. (1989). Telepathology: Long-distance diagnosis. *American Journal of Clinical Pathology, 91*, S32-S42.

Weller, E. L., & Mahoney, G. J. (1983). Comparison of oral and total communication modalities on the language training of young mentally handicapped children. *Education and Training of the Mentally Retarded, 18*, 103–110.

Westman, M. J., & Broen, P. A. (1989). Pre-school screening for significant articulation errors. *Language, Speech, and Hearing Services in Schools, 20*, 139–148.

Whittaker, S., & O'Conaill, B. (1997). The role of vision in face to face and mediated communication. In K. E. Finn, A. J. Sellen, & S. J. Wilbur (Eds.), *Video mediated communication.* Hillsdale: Lawrence Erlbaum Associates.

Whitten, P., & Collins, B. (1997). The diffusion of telemedicine. *Science Communication, 19*, 21–40.

Williams, E. (1977). Experimental comparisons of face to face and mediated communication. *Psychological Bulletin, 84*, 963–976.

Wilson, L., Gill, R. A., Sharp, I. F., & Heitmann, S. A. (2000). The hospital without walls: A home telecare system incorporating vital signs monitoring. *Telemedicine Journal, 6*, 1–8.

Wittson, C. L., & Benschoter, R. (1972). Two-way television: Helping the medical centre reach out. *American Journal of Psychiatry, 129*, 136–139.

Wright, P. C., & Monk, A. (1989). Evaluation for design. In A. Sutcliffe & L. Macaulay (Eds.), *People and computers V* (pp. 345–358). Cambridge, UK: Cambridge University Press.

Zarate, C. A., Weinsrock, L., Cukor, P., Morabito, C., Leahy, L., Burns, C., & Baer, L. (1997). Applicability of telemedicine for assessing patients with schizophrenia: Acceptance and reliability. *Journal of Clinical Psychiatry, 58.*

Appendix A

Specific Details of the Impairments of the Children in the TeachSpeech Group

Child	Description of Problem
A	Mild receptive and expressive language delay, short attention
B	Mild receptive and expressive language delay, weak attention and short term memory
C	Semantic and pragmatic difficulties, possible atypical autism and mild receptive and expressive language delay
D	Moderate phonological disorder
E	Global delay with good memory and delayed echolalia
F	Moderate receptive language delay, severe expressive language delay, phonology disorder
G	Moderate receptive and expressive delay
H	Moderate receptive and severe expressive language delay - elective mute
I	Duchennes Muscular Dystrophy

Appendix B

Specific Details of the Impairments of the Children in the Comparison Group

Child	Description of Problem
J	Language delay (receptive and expressive)
K	Delay in comprehension and expression. Listening and attention weak
L	Pragmatic difficulties
M	Language disorder, word finding difficulties
N	Fronting / cluster reduction and final consonant deletion / gliding
O	Language delay - deaf parents
P	Language delay and behavioural problems
Q	Language and phonology delay leading to difficult behaviour
R	Moderate learning difficulties
S	Language and phonology delay
T	Phonology problems
U	Language delay - bilingual
V	Word finding difficulties, expressive language delay
W	Phonology disorder
X	Phonology and language
Y	Speech and language delay

Appendix C

Cost Analysis as a Pilot Project

Fixed Costs

* Cost of therapist

Several steps were taken to calculate the cost of the therapists' time on a per minute basis. The final result was based on an annual salary of £21,743 (spine scale 30)[1] which was divided by 52 (the weeks in the year) and then by 40 (number of hours in a working week). Finally, this figure was divided by 60 (number of minutes in an hour). The cost of the therapists time per minutes was calculated to be 17 pence per minute.

	TeachSpeech	Comparison
Cost of therapist in pence per minute	17	17
Amount of therapists time in minutes	57.3	61.0
Number of months	36	36
Number of children	9	9
Total	£3,156	£3,360

[1] Data supplied by the Information Team at the Royal College of Speech and Language Therapy

* Equipment

	TeachSpeech	Comparison
BT VC6000[2] videoconferencing units	£7,950	N/A
Number of units	5	N/A
Total	£39,750	N/A

* Start up costs - These costs include installation and connection of appropriate ISDN lines.

	TeachSpeech	Comparison
Start up charges	£199	N/A
Number of units	5	N/A
Total	£995	N/A

Direct Variable Costs

* Travel cost

Both the Comparison and TeachSpeech therapists receive 39.8 pence per mile in expenses[3]. It should be noted that the expenses are to cover both petrol and veichle degradation. An assumption held is that the therapist travel at an average of 30 miles per hour. Therefore it was assumed that the TeachSpeech therapist travelled 3 miles, based on the 6 minutes of travel time utilised per month. Likewise, it was assumed the NHS therapist travelled 8.35 miles, based on the 16.7 minutes of travel time utilised per month.

	TeachSpeech	Comparison
Number of miles travelled on average per month	3	8.35
Number of months	36	36
Number of children	9	9
Cost of travel per mile	39.8	39.8
Total	£387	£1077

[3]Information received from the NHS Salisbury District Hospital

* Line rental and calls

	TeachSpeech	Comparison
Average line rental per quarter	£380	N/A
Number of quarters in the year	4	N/A
Number of years	3	N/A
Number of videoconferencing units	5	N/A
Total	£22,800	N/A

Indirect Variable Costs

* Insurance

	TeachSpeech	Comparison
Insurance per year per unit	£200	N/A
Number of years	3	N/A
Number of videoconferencing units	5	N/A
Total	£3,000	N/A

Total Overall Cost

TeachSpeech = £70,088

Comparison = £4,437

Appendix D

Cost Analysis as a Service

To calculate the costs of both models of therapy as a service certain key assumptions need to be made:

- There are forty children.
- There are four schools.
- All the schools are situated 30 miles from the SLTs base.
- A return journey to each school takes 80 minutes.
- The time period is five years.

Fixed Costs

* Cost of therapist

Firstly, the travel times for each model of therapy needed to be adjusted. To do this the number of times per month the respective SLTs visited a child was determined from the log sheets. This was done for both models of therapy. It was found that the SLT in the TeachSpeech group visited the a child 0.28 times per month and that the SLT in the comparison group visited each child 0.64 times per month. This was then multiplied by the amount of time it would take the SLT to make a round trip to the school and back (80 minutes). It was calculated that the TeachSpeech SLT would be travelling for a total of 11 minutes per month per child and that the comparison group SLT would be travelling for an average of 26 minutes per month per child. This amount of travel time is added to the total amount of time (minus the original travel time) to obtain the total amount of time per child the therapist spends on the child.

Secondly, the salary of the SLT needed to be calculated. The same format as in Appendix A is used below. The calculation is based on an annual salary of £21,743 (spine scale 30) this is divided by 52 (the number of weeks in a year) and by 40 (the number of

hours in a working week) and finally by 60 (the number of minutes in an hour). As in Appendix A this equals 17 pence per minute.

	TeachSpeech	Comparison
Cost of therapist in pence per minute	17	17
Amount of therapists time in minutes	68.3	87.0
Number of months	60	60
Number of children	40	40
Total	£27,866	£35,496

* Equipment

	TeachSpeech	Comparison
BT VC6000 videoconferencing units	£7,950	N/A
Number of units	5	N/A
Total	£39,750	N/A

* Start up costs

	TeachSpeech	Comparison
Start up charges	£199	N/A
Number of units	5	N/A
Total	£995	N/A

Direct Variable Costs

* Travel cost

The travel costs are calculated by finding the number of times a therapist would travel to a school to see a child over the five year period and multiplying it the total number of miles a round trip to a school would be. This gives the total number of miles travelled by the therapist. From this, the cost in expenses paid to the therapist can be calculated.

	TeachSpeech	Comparison
Number of times a therapist travels to see a child per month	0.28	0.64
Months	60	60
Children	40	40
Total times the therapist visits the school	672	1536

	TeachSpeech	Comparision
The total number of times a therapist travels to see a child over a five year period	672	1536
Miles	60	60
Cost per mile in pence	39.8	39.8
Total cost	£16,047	£36,680

* Line rental and calls

	TeachSpeech	Comparision
Average line rental per quarter	£380	N/A
Number of quarters in the year	4	N/A
Number of years	5	N/A
Number of videoconferencing units	5	N/A
Total	£38,800	N/A

Indirect Variable Costs

* Insurance

	TeachSpeech	Comparision
Insurance per year per unit	£200	N/A
Number of years	5	N/A
Number of videoconferencing units	5	N/A
Total	£5,000	N/A

Total Overall Cost

TeachSpeech = £127,658

Comparison = £72,176

Appendix E

Analysis of Questionnaires Given to TeachSpeech SLT

The response the participant gave the first time the questionnaire was filled in has been marked with a F. The response the participant gave the second time the questionnaire was filled in has been marked with a S. Some questions were not in the second questionnaire. These sections have been marked with a †. Questions that were in the second questionnaire only, are marked with a ‡.

PERSON SPECIFICATION FOR THE FIRST AND SECOND SET

Gender :	1 female
Age :	1 36 - 45

ABOUT THE TECHNOLOGY - AUDIO

Q1 How easy was it to hear and understand the remote participant, (in most cases this refers to Karen)?

Very easy	S
Fairly easy	F
Neither easy nor difficult	0
Fairly difficult	0
Very difficult	0

If you experienced any difficulty, please indicate these problems below.

F - The microphone supplied with the VC system cuts in and out at random, the tie microphones work much better.

Q2 How quick and easy was it to get your questions across to remote participant?

Very easy	FS
Fairly easy	0
Neither easy nor difficult	0
Fairly difficult	0
Very difficult	0

If you experienced any difficulty, please indicate the problems below.
No comments made

Q3 Did you experience any audio delay during the sessions?

Yes	FS
No	0
Unsure	0

Q4 In general how easy was it for you to make a verbal contribution to the sessions when you wanted?

Very easy	FS
Fairly easy	0
Neither easy nor difficult	0
Fairly difficult	0
Very difficult	0

Q5 How easy did you feel it was to explain your point, as compared to a situation when you are face to face?

Very easy	FS
Fairly easy	0
Neither easy nor difficult	0
Fairly difficult	0
Very difficult	0

† Q6 In general how often did you attempt to speak at the same time as another remote participant?

Never	0
Rarely	F
Not very frequently	0
Fairly frequently	0
Very frequently	0

† Q7 If this did occur, did you stop speaking and wait for the other person to resume speaking?

Never	0
Rarely	0
Not very frequently	0
Fairly frequently	0
Very frequently	F

ABOUT THE TECHNOLOGY - VIDEO

Q1 How would you rate the quality of the remote video image(s)?

Excellent - "TV Quality"	0
Very good	FS
Average	0
Fairly poor	0
Poor - "Grainy / Indistinct"	0

If you experienced any difficulty, please indicate the problems below.

F - No Comments Made

S - About 10% of the calls experience green and pink squares which reduce the picture quality. This represents a poor line and necessitates re-dialling.

Q2 During the session, how often did you look at remote participant?

Never	0
Rarely	0
Not very frequently	0
Fairly frequently	0
Very frequently	FS

Q3 Did you feel that the availability of a video image helped you to get your point across?

Never	0
Rarely	F
Not very frequently	0
Fairly frequently	0
Very frequently	FS

S - It focuses the conversation. People seem aware of the time limitations and costs involved, unlike during a school visit. Concentration required, perhaps due to the slight discrepancy in audio / visual signals.

Q4 Did you feel that the availability of a video image helped you to understand what remote participant was saying?

Never	0
Rarely	0
Not very frequently	0
Fairly frequently	F
Very frequently	S

F - Sometimes it made no difference

S - You see the emphasis and facial expression that goes with the description.

Q5 Did you feel that you were missing certain non-verbal communication signals made by the remote participants, e.g. gestures or facial expressions?

Never	FS
Rarely	0
Not very frequently	0
Fairly frequently	0
Very frequently	0

SOCIAL ASPECTS OF THE TECHNOLOGY

† Q1 Prior to this session, had you met remote participant before in a face to face context?

Yes	F
No	0
Unsure	0

Q2 Did you experience a "sense of presence" with the other participant?

Never	0
Rarely	0
Not very frequently	0
Fairly frequently	0
Very frequently	FS

S - Telephone us seems to suit only straightforward conversation - not complex information sharing.

Q3 Do you feel able to focus on the agenda of the session?

Yes	FS
No	0
Don't know	0

Q4 Do you feel you understood and take in what remote participant is saying?

Yes	FS
No	0
Don't know	0

If yes, do you feel that the session is clearly reflected in the work you later completed?

Yes	FS
No	0
Don't know	0

Q5 Do you feel free to chat informally during the sessions?

Yes	FS
No	0
Don't know	0

Q6 In general how satisfied were you with the outcome of the sessions?

Very satisfied	FS
Fairly satisfied	0
Neither satisfied or dissatisfied	0
Fairly dissatisfied	0
Very dissatisfied	0

Please explain your answer further.

F- The information that was given was concise and focused It eliminates some distractions

S - It offers enhanced communication. It helps to establish a working relationship and it more clearly communicates the message.

Q7 Can you think of any advantages that this technology provided as compared with face to face sessions?

Yes	FS
No	0

If yes, please expand.

F- Helps to focus on the topic. During Face to Face visits you can be distracted by peripheral noise / voices. All of the materials needed are at hand. It is difficult to find a quiet place to talk. Cheaper in the long run. Therapist will be able to see more children.

S - The therapist will be able to advise about a child more frequently. Delivery of training to teachers, parents and ESAs.

Q8 Can you think of any disadvantages that this technology provided as compared with face to face sessions?

Yes	S
No	F

If yes, please expand:

F - Quality of image may not be good enough. Expense to install. Apprehensive that videoconferencing will be a complete alternative

S - The child needs to be assessed face to face and observed in the school. The range of the camera and limitations of ISDN lines limit the observation use.

† Q9 Do you feel this technology has future value to your profession?

Yes	F
No	0
Don't know	0

† Q10 Do you feel this technology has future value to other professions?

Yes	F
No	0
Don't know	0

If yes, please state which professions.

Educational Psychologists, Support Teachers, Teachers for Hearing Impaired, Physio and Occupational Therapists

† Q11 Have you used or been aware of multimedia communication tools that provide a speech and language service?

Yes	0
No	F

If yes, please give details.

No Comments Made

PRIOR COMPUTER EXPERIENCE

† Q1 How much experience have you had of using computers?

Very experienced	0
"I use computers nearly everyday"	
Quite Experienced	F
"I have done a fair amount of work using computers, but there are still lot of things I don't know"	
Moderately experienced	0
"I use them sometimes but my skills are limited"	
Not very experienced	0
"I rarely use computers"	
No experience	0
"I've not really used them at all and don't know how to"	

† Q2 Have you ever used E-mail in your work environment or for personal use?

Yes	0
No	F
Don't know	0

COMPARISON WITH EARLIER EXPERIENCE

‡Q1. Do you feel there are any changes in the way that you communicate, over the video-conferencing link, as compared with when you first started?

Yes	S
No	0
Don't know	0

Please explain your answer.

S - I feel more confident using the link. I've noticed even sceptical ESA's, teachers and parents become more confident and comfortable using it time.

‡Q2 Do you feel there are any changes in your working environment since you started to use the link?

Yes	0
No	0
Don't know	S

Please explain your answer.

No Comments Made

‡Q3 Do you feel that your working practice has changed in any way?

Yes	0
No	S
Don't know	0

Please explain your answer.

No Comments Made

Appendix F

Analysis of Questionnaires Given to the TeachSpeech ESAs

The responses the participants gave the first time the questionnaire was filled in has been marked with a F. The responses the participants gave the second time the questionnaire was filled in has been marked with a S. Some questions were not in the second questionnaire. These sections have been marked with a †. Questions that were in the second questionnaire only, are marked with a ‡.

PERSON SPECIFICATION FIRST SET
Gender : 5 females
Age : 1 20 - 30, 3 31 - 40, 1 41 - 50

PERSON SPECIFICATION SECOND SET
Gender : 5 females
Age : 1 20 - 30, 4 31 - 40, 1 41 - 50
Note one person had not filled in the questionnaire as part of the first set so their results were added to the first (f) set.

ABOUT THE TECHNOLOGY - AUDIO

Q1 How easy was it to hear and understand the remote participant, (in most cases this refers to Karen)?

	F	S
Very easy	2	2
Fairly easy	4	2
Neither easy nor difficult	0	0
Fairly difficult	0	0
Very difficult	0	0

If you experienced any difficulty, please indicate these problems below.

F - At one point problems due to BT work (they were unaware of link) Echoing can occur, but can be solved by using volume button

S - Can just use volume control

Q2 How quick and easy was it to get your questions across to remote participant?

	F	S
Very easy	3	3
Fairly easy	3	1
Neither easy nor difficult	0	0
Fairly difficult	0	0
Very difficult	0	0

If you experienced any difficulty, please indicate the problems below.

No comments made

Q3 Did you experience any audio delay during the sessions?

	F	S
Yes	5	3
No	0	0
Unsure	1	1

F - Only when making connection at the beginning of the session. 2-3 second delay

S - At times, but doesn't alter conversation. Its annoying as you interrupt them.

Q4 In general how easy was it for you to make a verbal contribution to the sessions when you wanted?

	F	S
Very easy	3	2
Fairly easy	3	2
Neither easy nor difficult	0	0
Fairly difficult	0	0
Very difficult	0	0

Q5 How easy did you feel it was to explain your point, as compared to a situation when you are face to face?

	F	S
Very easy	2	3
Fairly easy	3	0
Neither easy nor difficult	1	1
Fairly difficult	0	0
Very difficult	0	0

† Q6 In general how often did you attempt to speak at the same time as another remote participant?

Never	0
Rarely	3
Not very frequently	2
Fairly frequently	0
Very frequently	0

After a few sessions you adapt.

† Q7 If this did occur, did you stop speaking and wait for the other person to resume speaking?

Never	0
Rarely	2
Not very frequently	0
Fairly frequently	2
Very frequently	1

ABOUT THE TECHNOLOGY - VIDEO

Q1 How would you rate the quality of the remote video image(s)?

	F	S
Excellent - "TV Quality"	1	1
Very good	4	2
Average	1	1
Fairly poor	0	0
Poor - "Grainy / Indistinct"	0	0

If you experienced any difficulty, please indicate the problems below.

F - Picture is not always as clear as a TV image. Visual aids can be seen clearly when held near the screen

S - Green squares can occur and can't show pictures as well. When the system hasn't been used for a while sometimes get green squares.

Q2 During the session, how often did you look at remote participant?

	F	S
Never	0	0
Rarely	0	0
Not very frequently	0	0
Fairly frequently	2	0
Very frequently	4	4

Q3 Did you feel that the availability of a video image helped you to get your point across?

	F	S
Never	0	0
Rarely	0	0
Not very frequently	0	0
Fairly frequently	3	2
Very frequently	2	2

F - 1 person gave no answer

Q4 Did you feel that the availability of a video image helped you to understand what remote participant was saying?

	F	S
Never	0	0
Rarely	0	0
Not very frequently	0	0
Fairly frequently	2	1
Very frequently	3	3

F - 1 person gave no answer. Visual aids and gestures help you understand.

Q5 Did you feel that you were missing certain non-verbal communication signals made by the remote participants, e.g. gestures or facial expressions?

	F	S
Never	3	3
Rarely	3	1
Not very frequently	0	0
Fairly frequently	0	0
Very frequently	0	0

F - All gestures can be seen.

SOCIAL ASPECTS OF THE TECHNOLOGY

† Q1 Prior to this session, had you met remote participant before in a face to face context?

Yes	3
No	2
Unsure	0

Q2 Did you experience a "sense of presence" with the other participant?

	F	S
Never	0	0
Rarely	0	0
Not very frequently	0	0
Fairly frequently	2	0
Very frequently	3	4

F - 1 person gave no answer. You often feel in the same room as the other person.

S - You often feel in the same room as the other person.

Q3 Did you feel able to focus on the agenda of the session?

	F	S
Yes	6	4
No	0	0
Unsure	0	0

Q4 Did you feel you understood and took in what remote participant was saying?

	F	S
Yes	6	4
No	0	0
Unsure	0	0

If yes, do you feel that the session was clearly reflected in the work you later completed?

	F	S
Yes	6	4
No	0	0
Unsure	0	0

F - Helpful as can make use you are following the correct guidelines for that child

Q5 Did you feel free to chat informally during the sessions?

	F	S
Yes	6	4
No	0	0
Unsure	0	0

S - But aware of costs.

Q6 In general how satisfied were you with the outcome of the sessions?

	F	S
Very satisfied	4	3
Fairly satisfied	2	0
Neither satisfied or dissatisfied	0	1
Fairly dissatisfied	0	0
Very dissatisfied	0	0

Please explain your answer further.

F- Sessions focus my time and allow me to spend more time with the child. It felt as if I was talking to the person face to face

S - Karen 'shows' us what to do.

Q7 Can you think of any advantages that this technology provided as compared with face to face sessions?

	F	S
Yes	5	4
No	1	0

If yes, please expand.
F - Helps the child in the best possible way with a minimum time outlay. There is a quick response time to any problems that occur. Time travelling reduced. Both people who answered yes above, ticked both boxes in the expand section. Cheaper in the long run. Therapist will be able to see more children
S - Time saving.

Q8 Can you think of any disadvantages that this technology provided as compared with face to face sessions?

	F	S
Yes	4	3
No	2	1

If yes, please expand:

	F	S
Quality of image may not be good enough	1	0
Expense to install	2	2
Apprehensive that videoconferencing will be a complete alternative	2	2
Other - please give details	0	0

† Q9 Do you feel this technology has future value to your profession?

Yes	5
No	0
Don't know	0

† Q10 Do you feel this technology has future value to other professions?

Yes	4
No	0
Don't know	1

If yes, please state which professions.

F - Educational Psychologists, Behavioural Support, Fire brigade, Nursing, Police

† Q11 Have you used or been aware of multimedia communication tools that provide a speech and language service?

Yes	1
No	4

If yes, please give details.

No Comment made.

PRIOR COMPUTER EXPERIENCE

† Q1 How much experience have you had of using computers?

Very experienced	0
"I use computers nearly everyday"	
Quite Experienced	2
"I have done a fair amount of work using computers, but there are still lot of things I don't know"	
Moderately experienced	2
"I use them sometimes but my skills are limited"	
Not very experienced	0
"I rarely use computers"	
No experience	1
"I've not really used them at all and don't know how to"	

† Q2 Have you ever used E-mail in your work environment or for personal use?

Yes	0
No	5

COMPARISON WITH EARLIER EXPERIENCE

‡ Q1. Do you feel there are any changes in the way that you communicate, over the video-conferencing link, as compared with when you first started?

Yes	3
No	1
Don't know	0

Please explain your answer.

More relaxed as have now met face to face. Always found it easy. Feels more like face to face now. We are more at ease with talking into the microphone.

‡ Q2 Do you feel there are any changes in your working environment since you started to use the link?

Yes	1
No	2
Don't know	1

Please explain your answer.

Staff more interested. Had a meeting to show them how it works.

‡ Q3 Do you feel that your working practice has changed in any way?

Yes	2
No	1
Don't know	1

Please explain your answer.

People are more at ease with TeachSpeech Project. I am now much more skilled in the area of my work, this reflects when I work with other children too.

Appendix G

Analysis of Questionnaires Given to the TeachSpeech Parents

The response the participant gave the first time the questionnaire was filled in has been marked with a F. The response the participant gave the second time the questionnaire was filled in has been marked with a S. Some questions were not in the second questionnaire. These sections have been marked with a †. Questions that were in the second questionnaire only, are marked with a ‡.

PERSON SPECIFICATION FOR THE FIRST SET Gender: 8 females
Age: 5 21 - 30, 2 41 - 50, 1 unknown
Occupation: Housewife x 4 / Textile retailer / Cleaner / Shop Assistant / Tax Consultant

PERSON SPECIFICATION FOR THE SECOND SET
Gender: 5 females
Age: 2 21 - 30, 1 31 - 40, 2 41 - 50
Occupation: Housewife x 4 / Shop Assistant

SOCIAL ASPECTS

Q1 How satisfied are you with the teaching that your child is being given from the ESA?

	F	S
Very satisfied	3	5
Fairly satisfied	3	0
Neither satisfied nor dissatisfied	1	0
Fairly dissatisfied	0	0
Very dissatisfied	0	0

If you wish, you may explain your answer further.

Support Limited. Lack of Feedback

Q2 How satisfied are you with the progress of your child in relation to the help that your child is receiving?

	F	S
Very satisfied	6	4
Fairly satisfied	1	1
Neither satisfied nor dissatisfied	1	0
Fairly dissatisfied	0	0
Very dissatisfied	0	0

If you wish, you may explain your answer further.

ESA more confident as receiving weekly support, in turn this helps child.

Q3 Do you have any reservations with the Speech and Language Therapist using a video conferencing system to give support to the ESA?

	F	S
Yes	0	0
No	8	5

If you wish, you may explain your answer further.

Closer links with ESA. Allows therapist to give advice to ESA's who are located in a wider area.

Q4 Do you see any advantages of using such technology to give support?

	F	S
Yes	7	3
No	1	2

If you wish, you may explain your answer further.
F -Child does not need to travel to SLT. Other children will benefit from a more experienced ESA . ESA is able to check / show work to SLT which is helpful. My son is happier at mainstream school and this way he still has contact with the speech and language therapist. More children can be given help on a more frequent basis. The support given to my child by the video conferencing system enables him to stay in an environment he is comfortable in and saves taking time out to have appointments.
S - It enables the child to receive more therapy. Saves travelling time. The school is more aware of the child's problems.

Q5 Imagine that the Speech and Language therapist is giving support to the ESA in person rather than across a video link. Do you feel that there would be any difference in the quality of care your child is receives?

	F	S
Yes	4	2
No	4	3

If you wish, you may explain your answer further.
F -If face to face, less sessions would occur and confidence would drop. There would be fewer distractions for the child (*). I don't think anything beats personal contact although the quality of the link will be excellent.
S - He would not be receiving any speech therapy if it wasn't for TeachSpeech. They are aware of the cost and conversation is stifled.

Q6 Do you feel there might be any disadvantages with the videoconferencing technology?

	F	S
Yes	3	3
No	5	1

If yes, please expand.

	F	S
Quality of image may not be good enough	0	2
Expense	2	2
Apprehensive that videoconferencing may become a complete alternative	0	2

Other - please give details

F - Child distracted (*)

S - We have had a lot of problems with BT lines.

† Q7 Before TeachSpeech had you heard of any videoconferencing systems that provided a speech and language service?

Yes	7
No	1

If you wish, you may explain your answer further.

No comments made

PRIOR COMPUTER EXPERIENCE

† Q1 How much experience have you had of using computers?

Very experienced "I use computers nearly everyday"	1
Quite Experienced "I have done a fair amount of work using computers, but there are still lot of things I don't know"	1
Moderately experienced "I use them sometimes but my skills are limited"	0
Not very experienced "I rarely use computers"	2
No experience "I've not really used them at all and don't know how to"	4

† Q2 Before your child started TeachSpeech had you ever seen how a videoconferencing system operates?

Yes	1
No	7

If you wish, you may explain your answer further.

No comments made

If there are any comments you wish to make that this questionnaire has not addressed please do so:

F - Thank you for all your help, my child has become a little chat-a-box. Would like to have TeachSpeech for my child though out his school life. Having TeachSpeech feels like the school is interested and my child is being helped appropriately. It is early days yet to notice what effect this is having. TeachSpeech is good for continuity - my child sees the same therapist. My child is receiving much more help than he ever had when he was at the hospital.

N.B. (*) This parent seems to think that there child is receiving help from the SLT directly across the video conferencing link.

Appendix H

Analysis of Questionnaire Given to the Comparison SLTs

PERSON SPECIFICATION

Gender: 2 females

Age: 1 21-30, 1 31-40

ABOUT YOUR WORK

Q1 Approximately what is your case load?

0 - 75 children	0
76 - 150 children	1
151 - 225 children	1
226 - 300 children	0
Over 301 children	0

Q2 Approximately how much time do you spend in sessions with children in a week, (this refers to your entire case load)?

Under 5 hours per week	0
6 - 10 hours per week	0
11 - 15 hours per week	0
16 - 20 hours per week	1
Over 21 hours per week	1

Q3 On what basis do you decide whether you see a child for a session or not?

Severity of Impairment	1
Type of Impairment	1
Whether the child has access to an Educational Support Assistant (ESA) or not	1
How willing the child is to travel	0
How far it would be for you or the child to travel	0

All children are seen who are referred for an initial assessment and then decide depending on there 'problem' what sort of treatment they'll receive.

Q4 Approximately how much of your time is taken up doing administration work?

Under 5 hours per week	0
6 - 10 hours per week	1
11 - 15 hours per week	1
16 - 20 hours per week	0
Over 21 hours per week	0

Q5 Approximately how often do you see an ESA (any ESA)?

Rarely - Less than once a month	1
Seldom - About once a month	0
Occasionally - About once every two weeks	0
Frequently - Once a week	0
Very frequently - More then once a week	1

Each time I see a child I see an ESA.

Q6 Do you think it would be useful to see an ESA more?

Yes	1
No	1

If you wish, please explain your answer further.
Clinical commitments mean that this isn't possible.

Q7 If you would like to spend more time with the Educational Support Assistant, approximately how much additional time would be ideal?
1 hour per month per child.

Q8 Approximately how much time each week do you spend travelling, to see ESA?

Less than 1 hour per week	1
2 - 3 hours per week	0
4 - 5 hours per week	1
6 - 7 hours per week	0
Over 7 hours per week	0

THE SESSIONS WITH ESA

Q1 Do you feel able to focus on the agenda of a session?

Very easy	1
Fairly easy	0
Neither easy nor difficult	0
Fairly difficult	0
Very difficult	0

1 N/A

Q2 Do you feel free to chat informally during the sessions?

Yes	1
No	0
Don't know	0

1 N/A

Q3 In general how satisfied are you with the outcomes of the sessions?

Very satisfied	1
Fairly satisfied	0
Neither satisfied or dissatisfied	0
Fairly dissatisfied	0
Very dissatisfied	0

1 N/A

If you wish, please explain your answer further

ABOUT THE VIDEOCONFERENCING TECHNOLOGY

Q1 Do you feel this technology has value to your profession?

Yes	1
No	0
Don't know	1

If you wish please explain your answer further

An ESA is seen when assessing or providing therapy for a child, this is not possible via videoconferencing.

Q2 Do you feel that this approach would be useful in your work setting for others to provide input to you?

Yes	0
No	1
Don't know	1

If you wish please explain your answer further

Generally enough work is left for the ESA, no calls have been received re therapy. Any questions are saved until the school is visited

Q3 Do you feel that this approach would be useful for you to provide input to others?

Yes	1
No	1
Don't know	1

(One therapist answered both No and Don't Know)

If you wish please explain your answer further.

It would save on travelling time, but would lose contact with schools.

Q4 Can you think of any advantages that this technology provides as compared with face to face sessions?

Yes	2
No	0

If yes, please expand.

Cheaper in the long run	0
Therapist will be able to see the ESA more	1
Therapist will be able to see more children	0
Other	0

The therapist would have to trust that the ESA was doing / providing the therapy appropriately.

Q5 Can you think of any disadvantages that this technology provides as compared with face to face sessions?

Yes	1
No	0

If yes, please expand:

Quality of image may not be good enough	1
Expensive to install	2
Apprehensive that videoconferencing will become a complete alternative	2

Other - please give details

Important to build rapport with children using face to face contact. Videoconferencing would not allow a good relationship to be established

PRIOR COMPUTER EXPERIENCE

Q1 How much experience have you had of using computers?

Very experienced	0
"I use computers nearly everyday"	
Quite Experienced	0
"I have done a fair amount of work using computers, but there are still lot of things I don't know"	
Moderately experienced	1
"I use them sometimes but my skills are limited"	
Not very experienced	1
"I rarely use computers"	
No experience	0
"I've not really used them at all and don't know how to"	

Q2 Have you ever used E-mail in your work environment or for personal use?

Yes	0
No	2

Appendix I

Analysis of Questionnaires given to the Comparison Parents

PERSON SPECIFICATION

Gender: 9 females

Age: 1 21 - 30, 4 31 - 40, 1 41 - 50, 3 unknown

Occupation: Housewife x 2 / Telephonist / Sales Administrator / Staff Nurse / Water Company Inspector / 3 Unknown.

Q1 How satisfied are you with the support that your child is being given from the speech and language therapist?

Very satisfied	3
Fairly satisfied	4
Neither satisfied nor dissatisfied	1
Fairly dissatisfied	0
Very dissatisfied	1

If you wish, you may explain your answer further. James attended an I CAN nursery. Waiting list long. Never sees Speech and Language therapist.

Q2 How satisfied are you with the progress of your child in relation to the help that your child is receiving?

Very satisfied	3
Fairly satisfied	3
Neither satisfied nor dissatisfied	1
Fairly dissatisfied	0
Very dissatisfied	1

If you wish, you may explain your answer further.
My child received therapy from 2 till 5 years. His speech is now entirely normal even though he could not speak at all before he was 2 1/2. Progress not improved since has hardly any support (3hours in 3 1/2 years).

THE ALTERNATIVE MODEL

Q1 Would you have any reservations with the Speech and Language Therapist using a video conferencing system to give support to the ESA?

Yes	2
No	6

If you wish, you may explain your answer further.
Only a SLT supports my son, he gets no ESA time.

Q2 Do you see any advantages of using such technology to give support?

Yes	7
No	1

If you wish, you may explain your answer further.
For other children who need a lot of speech and language therapy. I should imagine that the speech therapist could use her time more effectively. Continuity, same goals and plan of attack. My child's problem was difficulty in one to one communication so I don't think it would help.

Q3 Imagine that the Speech and Language therapist is giving support to the ESA over the video link rather than in person. Do you feel that there would be a decline in the quality of care your child is receives?

Yes	2
No	6

If you wish, you may explain your answer further.
The children may feel distracted by seeing themselves on television. I've been told that the video link works very well. The ESA has not received the same training as the SLT so I think there would be a difference.

Q4 Do you feel there might be any disadvantages with the videoconferencing technology?

Yes	6
No	3

If yes, please expand.

Quality of image may not be good enough	2
Expense	2
Apprehensive that videoconferencing may become a complete alternative	3

Other - please give details
Children prefer face to face contact. My child received 40mins of one to one time, I can't imagine the ESA being able to do this and they are not as well qualified

Q5 Have you heard of any videoconferencing systems that provided a speech and language service?

No	4
Yes	5

If you wish, you may explain your answer further.
My son was involved in the launch of the I CAN nursery. It was introduced in my sons school after he had finished therapy. A friends child had the use of videoconferencing.

PRIOR COMPUTER EXPERIENCE

Q1 How much experience have you had of using computers?

Very experienced	3
"I use computers nearly everyday"	
Quite Experienced	2
"I have done a fair amount of work using computers, but there are still lot of things I don't know"	
Moderately experienced	1
"I use them sometimes but my skills are limited"	
Not very experienced	2
"I rarely use computers"	
No experience	1
"I've not really used them at all and don't know how to"	

Q2 Have you ever seen how a videoconferencing system operates?

Yes	1
No	8

If you wish, you may explain your answer further.

I would like to see one though.

If there are any comments you wish to make that this questionnaire has not addressed please do so:

I would love to hear the results and see if the videoconference system works.

Appendix J

Professional Evaluation Questionnaire

THE PROFESSIONS

Clinical Psychologist	2
Computer Science Lecturer	1
ESA	2
Head teacher	1
Information Technology	1
Occupational Therapist	1
Parent (Auxiliary nurse)	1
SALTA	11
Senior Education Officer	1
Special Educational Needs	1
Speech and Language Teacher	5
Speech and Language Therapist	23
Teacher	12
Total	62

Employing Authority	
Education	23
Health	30
Social Services	3
Other	4

Work Location	
Rural	11
Town	21
Other	9
Both	16
No Answer	5

COMPUTER EXPERIENCE

Q1 How much experience have you had of using computers?

Very experienced	4
Experienced	3
Moderately experienced	20
Limited experience	31
No experience	5

Q2 Have you ever used E-mail in your work environment or for personal use?

Yes	23
No	40

Q3 If Yes, to what extent have you used E-mail?

Regularly (more than twice a week)	9
Occasionally (less than twice a week)	13

Q4 Prior to this session, have you ever used any of the following types of multimedia communication tools (software)? You may tick more than one box.

Audio	43
Video	40
Whiteboard	5
(shared workspace with some graphics)	
Text tool	2
(shared workspace for text only)	
None of the above	15

Q5 If applicable, please indicate how experienced you are with the multimedia communication tools mentioned in Q4. You may tick more than one box.

	Audio	Video	Whiteboard	Text tool
Very experienced	6	5	1	1
Experienced	14	12	1	1
Limited experience	23	28	8	3

PROIR TO THE PRESENTATION

Q6 Have you used or been aware of multimedia communication tools that provide a speech and language service?

Yes	36
No	25

Details:

Debbie Osborn	4
Article in Therapy Weekly	2
Gave temp. support in speech and language unit	1
I CAN link in schools	7
Video Conference centre	2
Read in professional press	1
Computer games for specific SLT	2
Alex Hall gave a talk	1

Q7 Do you feel this technology has future value to your profession?

Yes	53
No	0
Don't know	7

Q8 Do you feel this technology has future value to other professions?

Yes	48
No	1
Don't know	11

Q9 Do you feel that this approach would be useful in your work setting for others to provide input to you?

Yes	47
No	1
Don't know	12

Q10 Do you feel that this approach would be useful for you to provide input to others?

Yes	45
No	3
Don't know	12

AFTER THE PRESENTATION

Q1 Do you feel that this approach would be useful in your work setting for others to provide input to you?

Yes	60
No	1
Don't know	1

Q2 Do you feel that this approach would be useful for you to provide input to others?

Yes	57
No	0
Don't know	6

Q3 Do you feel that multimedia communication technology would be useful to you?

Yes	60
No	0
Don't know	2

Reasons

For support services	5
Enable access to professional help	14
Sharing of Knowledge	3
Time saver	11
Linking with ESA / school regularly	5
Money saver	2
Link to a wide area	2
Closer contact with people	7
Quality is high	2
Fast / instant response to solve problems	5
Visual demonstration of techniques	2
Training Potential	3
Rural schools will benefit	2
Maximising minimal resources	1

Q4 Do you feel that this approach would be useful to another profession?

Yes	54
No	0
Don't know	3

Details : Educational Psychologist, Clinical Psychology, Behavioural Therapist, Occupational Therapist, Health and social services, Medical services, Psychiatry, Veterinary science, Teaching, Schools, Hearing Impaired, PASSIS, Special needs.

Q5 Do you see limitations / drawbacks to providing Speech and Language therapy input through this medium?

Yes -

Limitations	29
Drawbacks	7
Both	10

Reasons

Loss of social contact	11
Quality of Technology poor	1
Lack of space in schools	1
Still need face to face contact (for assessment)	10
Size of equipment	1
Clarity of Picture	3
Rapport with children lost	1
Not all facets of communication covered	1
Funding	10
Need confident / competent staff	8
Only limited number of children who benefit	3
Confidentiality	2
ESA should still be trained	1
Staff would need to be well trained	1
Its not a replacement	1
Showing the ESA techniques is much better	1
If you do it on the child and the ESA watches	1
Need positive support from schools	1

No - 13

Reasons

Because current therapy is limited it would be an addition to any therapy we can give	1
Not if set up properly with training	1
Work is more efficient and more effective	1
Providing it doesn't replace face to face contact	1
Increased level of support for children must be beneficial	1

Unsure -1

Many thanks for an informative visit.