THE EFFECT OF VIDEOTAPED FEEDBACK ON SPEECH DELIVERY AND SELF-EVALUATION BY ELEMENTARY SCHOOL CHILDREN

Robert J. D. Jones

A THESIS

in

The Department

of

Education

Presented in Partial Fulfillment of the Requirements for
the Degree of Master of Arts at
Sir George Williams University
Montreal, Canada

September, 1972

ABSTRACT

Robert J. D. Jones

THE EFFECT OF VIDEOTAPED FEEDBACK ON SPEECH DELIVERY AND SELF-EVALUATION BY ELEMENTARY SCHOOL CHILDREN

Seventy elementary school children, assigned to four treatment groups, delivered five speeches each with the inter-speech application of two independent variables: the presentation of a videotaped playback of the performance and the use of a self-evaluative checklist. The videotaped speeches were later evaluated by three judges. The hypotheses that all groups would show improvement in speech delivery performance over the five trials and that those groups receiving augmented feedback would show greater improvement were not statistically confirmed using the judges' evaluations as criteria. Using the subjects' self-evaluations it was hypothesized and confirmed that the subjects saw themselves improving significantly over five trials and that those receiving videotaped playback evaluated themselves lower, closer to the judges' evaluations, than did those not receiving such feedback. The subjects' oral evaluations and the researcher's observations provided further data on the effect of videotape self-confrontation.

PREFACE

This study grew out of a study proposed by Dr. Gary Boyd of Sir George Williams University entitled "A Study of the Optimization of Pupil and Student Behaviour Related to the Development of Knowledge Centres." Dr. Boyd's study proposed in part to examine the learner in interaction with the rich environment of the resource centre so as to determine whether optimal interaction patterns could be distinguished and so as to study the possibility of eliciting such patterns in resource-centre users. It was anticipated that the learning activities in a learner-centred environment would involve different processes than those in the instruction-oriented environment typical of the classroom. The type of learner who could best exploit the resource centre would be one capable of independent action, of defining his information needs, and of evaluating his own progress in achieving his goals with the minimum amount of other-person direction. It was this last element, that of learner self-evaluation, which became the prime concern of this investigator. Apart from possessing whatever information searching skills and strategies necessary to manoeuvre in the resource centre, it was felt that the learner would have to be able to monitor his own performance and to use this evaluation of his own progress to alter subsequent performances if he were to become an independent learner. It was proposed that the videotape recorder be used as a device to record learner behaviour for subsequent detailed study. At the same time it appeared that the videotape recorder could be used not only

as a recording device for researchers but also as a source of feedback to the learner.

A search of the literature indicated that although the videotape recorder had become a popular instrument in higher education, business training, sports and psychotherapy as a means of feedback, little research had been reported on its use in the elementary school where efforts were being made to develop learner-centred and independent-learner systems. Furthermore, it seemed that while many innovations had been adopted in educational systems, the field of evaluation had remained relatively untouched. Self-evaluation, while often lauded as superior to other-evaluation, was rarely put into practice except in a token manner. The evaluation that "counted" was teacher evaluation. Where self-evaluation was given priority, as in the free school, it seemed that little attempt had been made to determine or to develop the ability of the learner to assess his own performance systematically.

There appeared, then, to be a need for research on self-assessment at the elementary school level and for a study of the possibility of using videotaped feedback to enhance that self-evaluation. Because of the still poorly defined skills involved in effective resource-centre utilization and because of technical considerations involved in using the library-resource centre of the school selected for this research project, the original task of studying resource-centre utilization was not carried out in this study. In its place speech delivery, a psycho-motor skill capable of exploiting both the audio and video components of videotaped feedback, a skill with easily defined behavioural elements, and a skill which has traditionally depended heavily on other-evaluation, was chosen for

this exploratory study. The basic aim of this paper remained, however, a study of learner self-evaluation and the task chosen to study this process was seen primarily as a means of reaching this goal. It is hoped that the findings can be generalized beyond the specific task involved here to be of value to the original project which generated this study and to the general question of learner self-evaluation.

This study would not have been possible without the help and cooperation of several people and institutions. The study was financed by means of a grant from The Institut de Recherche Pédagogique du Ministère de l'Education of the Province of Quebec. Monies from the grant were used to purchase the video equipment and video tapes used in this study as well as to provide the investigator with an assistantship which allowed him to devote himself full-time to his studies and research.

Thanks are extended to the staff and pupils at Crestview Elementary School of the Protestant School Board of Greater St. Martin for their assistance and eagerness. Special thanks go to Mrs. Ann Johnston, principal, for her help and encouragement.

Dr. John MacDonald and Professor David Mitchell of the Department of Education at Sir George Williams University were both very helpful in providing critical questions to direct and structure the research in its early nebulous stages.

In particular, thanks are due to Dr. Gary Boyd, my thesis advisor. As recipient of the grant and director of the research project, he provided me with the opportunity to carry out this work. His very patient approach produced an ideal environment in which I could evolve my own ideas. This study would not have been possible without his

assistance, yet because of the non-obtrusive nature of his help, I can feel free to claim responsibility for its outcome in its successes as well as its failures.

Finally I would like to thank my wife, Evelyn, for her help in the long work of compiling the results, in library searches, in typing the first draft, and particularly for her patience. Thanks are also due to Mrs. Lee Smith and Mrs. Ellinor Williams for typing.

TABLE OF CONTENTS

PREFACE ii

TABLE OF CONTENTS vi

LIST OF TABLES viii

LIST OF FIGURES xi

INTRODUCTION

CHAPTER I. Scope of the Study 2

CHAPTER II. Need for the Study 5

CHAPTER III. Review of the Literature 14

EXPERIMENTAL DESIGN

CHAPTER IV. Hypotheses and Questions 36

CHAPTER V. Design of Study 39

CHAPTER VI. Description of Sample 50

CHAPTER VII. Instruments and their Characteristics 54

CHAPTER VIII. Data Analysis Plan 58

RESULTS AND CONCLUSIONS

CHAPTER IX Results 62

CHAPTER X. Discussion of Results 80

CHAPTER XI. Conclusion and Recommendations 86

LIST OF REFERENCES 92

APPENDICES

APPENDIX A Handout of General Information 98

APPENDIX B Data and Observation Forms 99

APPENDIX C Oral Instructions on How to Use the Speech Ruler 103

APPENDIX D The Speech Ruler 111

APPENDIX E Speech Checklist 115

APPENDIX F Oral Evaluation Analysis Form 118

APPENDIX G Video Recording and Playback Set-up 120

APPENDIX H Means of Judges' Evaluations, Self-evaluations, and Discrepancy Scores over Five Trials for the Four Treatment Groups 121

APPENDIX I Sex, I.Q., and Previous Speech Training Factors 127

LIST OF TABLES

Table		Page
1.	Designation of Treatment Groups According to the Distribution of Independent Variables	3
2.	Distribution of Subjects in Four Treatment Groups	52
3.	Analysis of Variance of Mean I.Q.s of Four Treatment Groups	53
4.	Means of Judges' Evaluations; Standard Deviations in Parentheses	63
5.	Analyses of Variance of Judges' Evaluations within each Treatment Group	65
6.	Multiple-classification Analysis of Variance of Judges' Evaluations of the Four Treatment Groups. Performance over Five Trials	66
7.	Means of Self-evaluations; Standard Deviations in Parentheses	67
8.	Analyses of Variance of Self-evaluations in Groups CHKL and VTR-CHKL	69
9.	Multiple-classification Analysis of Variance of Groups CHKL and VTR-CHKL's Self-evaluations over Five Trials	70
10.	Means of Discrepancy Scores; Standard Deviations in Parentheses	71
11.	Multiple-classification Analysis of Variance of the Discrepancy Scores of Groups CHKL and VTR-CHKL	72
12.	Mean Number of Oral Self-evaluative Statements by Subjects in Four Treatment Groups; Standard Deviations in Parentheses	73
13.	Analyses of Variance of the Number of Oral Self-evaluations among Four Treatment Groups	75

Table

Page

14. Means of Judges' Evaluations, Self-evaluations and Discrepancy Scores over Five Trials for the Four Treatment Groups 122

15. Means of Judges' Evaluations of Male and Female Subjects 127

16. Analyses of Variance of Judges' Evaluations within each Treatment Group: Trial X Sex Interaction 128

17. Means of Self-evaluations of Male and Female Subjects 129

18. Analyses of Variance of Self-evaluations within Treatment Groups: Trial X Sex Interaction 129

19. Means of Discrepancy Scores of Male and Female Subjects 130

20. Analyses of Variance of Discrepancy Scores within Treatment Groups: Trial X Sex Interaction 130

21. Means of Judges' Evaluations of Subjects Above and Below the Median I.Q. 131

22. Analyses of Variance of Judges' Evaluations within each Treatment Group: Trial X I.Q. Interaction 132

23. Means of Self-evaluation of Subjects Above and Below the Median I.Q. 133

24. Analyses of Variance of Self-Evaluations within each Group: Trial X I.Q. Interaction 133

25. Means of Discrepancy Scores of Subjects Above and Below the Median I.Q. 134

26. Analyses of Variance of Discrepancy Scores within each Group: Trial X I.Q. Interaction 134

27. Means of Judges' Evaluations of Subjects With and Without Previous Speech Training 135

28. Analyses of Variance of Judges' Evaluations within each Treatment Group: Trial X Training Interaction 136

Table

Page

29. Means of Self-evaluations of Subjects With and Without Previous Speech Training 137

30. Analyses of Variance of Subjects' Self-evaluations within each Group: Trial X Training Interaction 137

31. Means of Discrepancy Scores of Subjects With and Without Previous Speech Training 138

32. Analyses of Variance of Discrepancy Scores within each Group: Trial X Training Interaction 138

LIST OF FIGURES

Figure		Page
1.	Judges' Evaluations over Five Trials; Error Bars Showing Standard Deviations for VTR-CHKL and CNTL	63
2.	Means of Self-evaluations over Five Trials	68
3.	Means of Discrepancy Scores over Five Trials	71
4.	Mean Number of Oral Self-evaluative Statements in Four Treatment Groups	74

INTRODUCTION

CHAPTER I

SCOPE OF THE STUDY

This study involved the use of the videotape recorder as a mirror in which elementary school children could view their own speech performances. The effects of videotaped playback and of a self-evaluative checklist on pupil self-evaluation and on repeated performances of a speech delivery were studied. The following questions were asked: Do the addition of a videotaped playback of a speaking performance and the use of a self-evaluative checklist by the speakers lead to an improvement in subsequent performances? Do these variables have an effect on self-evaluation? How do elementary school children's self-evaluations differ from judges' evaluations? How do elementary school children react to self-confrontation via videotape?

The purpose of this study was 1) to exploit television as a learning aid rather than as an instructional aid; 2) to study the possibility of individualizing evaluation, that is, of developing learner-centred evaluation as a necessary corollary of learner-centred education; 3) to study the effect and effectiveness of videotaped playback on elementary school children.

While motivated and directed by all the needs expressed in the previous paragraph, this study limited its scope to the third, to the more specific question of the effectiveness of videotaped playback and checklist-aided self-evaluation in altering the speech delivery

performances of elementary school children, and to the differences between self- and other-evaluation. The experimental data were collected over a six week period using seventy elementary school children at the intermediate level who delivered five speeches each with inter-speech applications of two independent variables. The subjects were randomly assigned to four treatment groups designated and described according to the distribution of two independent variables: 1) the presentation of videotaped playback of a speech performance (VTR), and 2) the use of a self-evaluative checklist (CHKL), as outlined in the accompanying factor chart (Table 1). One group (VTR-CHKL) was exposed to both variables, another (CHKL) had only the use of the checklist, another (VTR) had only the videotaped playback, while the fourth (CNTL) had neither variable presented.

TABLE 1.--Designation of treatment groups according to the distribution of independent variables

Treatment Groups	Independent variable	
	Videotaped Playback	Self-evaluative Checklist
VTR - CHKL	Yes	Yes
CHKL	No	Yes
VTR	Yes	No
CNTL	No	No

Each subject's five speeches were evaluated later on videotape by three judges whose scores were averaged to yield a mean score for each speech. These mean scores plus the subjects' self-evaluations

in groups CHKL and VTR-CHKL were the dependent variables used in this study. Subjects also evaluated their performances orally and these evaluations, along with observations made by the experimenter, yielded data for a descriptive analysis of the effectiveness of the videotape recorder in promoting self-evaluation.

CHAPTER II

NEED FOR THE STUDY

This study was motivated by a concern for deficiencies perceived at several levels in the application of educational technology to the field of education. At the most general level, that concerning the role which educational technology should play in developing systems to support broad educational objectives, it was felt that too much emphasis had been placed on supporting traditional systems, with the result that educational technology faced the possibility of becoming a conservative force within education. At another level, it was felt that educational technology had rarely been applied to solving problems in evaluation of the learner's performance. At a more specific, but related, level, that concerning one possible role which the videotape recorder could play with young learners, a serious lack of information was found. This chapter will elaborate on these perceived deficiencies and indicate how it is felt this study may contribute to their elimination.

Much of what has been acclaimed as the "new media" and the "media revolution" in education has involved little more than the introduction of new hardware and software into traditional schools to support traditional, teacher-centred learning processes. Audio-visual specialists have been trying to place more equipment and materials into the hands of the teacher while educational innovators have tried to move away from the teacher as a controller and distributor of information

towards the student as director of his own learning activities. An examination of the learner's role in the teacher-with-instructional-aid mode indicates that the learner operates as a passive receiver and reactor to teacher-controlled and teacher-distributed information. While media used as an instructional aid might alter the teaching process and the teacher's activities it leaves the learner and his activities relatively unchanged.

With the development of instructional television, programmed instruction, computer-assisted instruction, and the systems-approach to instructional design, educational technology moved away from the use of media as an aid for the teacher to media as a replacement of the teacher. A distinction can be made between instructional aids, which support the classroom teacher who is the medium of instruction, and instructional media which carry the load of the instructional process with the teacher used as an aid or guide. The developers of such mediated systems of instruction are quick to point out that the teacher is not replaced but re-placed into a new role (Heinich, 1968). While these recent developments may have revolutionized the teaching process and the teacher's activities, the changes are not so radical from the learner's point of view. The live classroom teacher as controller and distributor of learning activities and information has been replaced by a mediated teacher, but the learner's activities still consist primarily of receiving and reacting to stimuli under some other person's control.

The "media revolution" has affected education, but outside the school. Television, for example, has radically changed the learner's access to information as well as the format of that information, not

in the classroom, but in the living room. The learner sitting in his living room deciding for himself whether or not to turn on the television set, deciding himself which of a variety of programmes to watch, and deciding himself how much involvement to invest in the programme is not the same learner, is not involved in the same learning process, as the learner sitting in a classroom receiving a programme selected by another person for that other person's objectives. Attempts to bring television into the classroom have altered television rather than the classroom. In the same way, the library book is different from the textbook, the filmstrip in the library-resource centre is different from the filmstrip shown in class.

The media revolution has not only increased the quantity and quality of the information available to the learner but has also resulted in an increased access to the means of production of mediated messages. The introduction of inexpensive and easily operated print duplicating, filming, and audio- and video-recording devices allows the individual and minority group to become the originators as well as receivers of communications. The videotape recorder used as a mirror in which to see and express one's self is not the same medium and does not generate the same responses as the same recorder used to record a commercial broadcast.

Educational technology's preoccupation with either supplying a support system for teacher-centred learning or with providing an alternative, the mediated teacher in system-controlled learning, has resulted in a lack of emphasis on the question of altering the learner's activities, of developing learning aids, and of designing alternative learning systems. The present study involved an attempt to explore

alternative learning processes, those generated by exploiting the unique capabilities of the videotape recorder as a medium which can give immediate access to audio-visual information about one's self. Television was put to use as a learning aid rather than an instructional aid, as an evaluation aid at the service of the learner rather than the teacher.

Evaluation of learning progress, while highly valued and considered essential in formal educational systems and in educational research, has remained relatively unaltered in a period which has seen the generation of alternative learning systems. Evaluation of the learner's progress in school learning has traditionally been the teacher's role. The teacher in teacher-centred learning systems not only designs and directs the learning process by supplying information and procedures to the learners, but also, at a teacher-determined point in the process, designs, administers and evaluates tests of the learners' performances. With the advent of media-based instructional systems the evaluative function shifted away from the teacher, except as a test administrator, to the system. From the learner's point of view, however, evaluation in both the teacher-centred and media-based systems is identical in that, like the instruction, it is other-controlled, other-generated and other-imposed. Even in those learning systems which allow the learner control of his learning experinces, final evaluation is very often in the hands of someone other than the learner. To the extent that this evaluation is used not merely as feedback to the learner but also as a sanctioned reward and punishment mechanism, to that extent the learner has lost the self-control desired in his learning experiences and has come under the control of the significant evaluators. For

a learner-centred and learner-controlled learning environment to function effectively the means for learner-directed self-evaluation must be provided.

Evaluation may be summative or formative (Bloom, Hastings and Madaus, 1971), system-oriented or learner-orinted. Traditionally educational evaluation has been summative, an event taking place upon completion of an instructional sequence, and system-oriented, an act whose objective it is to determine the status of the learner in the system. While summative and system-oriented evaluation may serve to determine the allocation of people and resources in an educational system and may be used to provide data for the revision of instructional programmes, formative and learner-oriented evaluation are needed by the learner to help him determine the degree of success or failure of his learning efforts. The more accurate and immediate that evaluative feedback is, the more likely it is that the learner will be able to use it to alter subsequent performances. Educational technology, while active in the area of instructional design, has done little in the way of proposing alternate evaluation designs. Perhaps because of their emphasis on instruction rather than on learning, educational technologists have tended to adopt summative evaluation. This thesis studied the effect that providing additional feedback to the performer had on self-evaluation and compared performer self-evaluation with other-evaluation.

The ease with which we can observe and measure a product rather than the process which produced it has resulted in most educational evaluation being based on the product of a performance. A product such as an answer on a test is more durable and observable than the ephemeral

and internal learning process which supposedly preceded and produced it. In evaluating a process by means of its product the assumption is made that a direct relationship exists between the product and the process. While this may be a safe assumption for many simple motor tasks, it becomes less certain with the complex and poorly defined behaviours commonly found in schooling. Emphasis on "what" is learned rather than on "how" it is learned is common in formal education and derives in part from this difficulty in observing processes as compared to products.

In many human activities, however, a product cannot be easily abstracted from the performance which produced it, or, quite often the performance is the product itself. Psychomotor skills such as those involved in oratory, dancing, and driving consist of performance-products so closely linked to a specific time and place that evaluation must take place either during the performance or afterwards through recollection. Evaluation of a psycho-motor skill in formal education is usually carried out by some judge other than the performer. When this evaluation is used in the formative mode as information about a performance fed back to a performer it should be recognized that this information, because of the complex processes of evaluation and communication involved in its transfer, has been drastically transformed from its original form, the performance. First the person judging a psycho-motor performance must, because of the quantity of information contained in the performance and because of its fleeting nature, attend to only small samples of the behaviour. The sampling will be influenced by the judge's value system which helps him determine which behaviour is worthwhile selecting for evaluation.

He will, in a sense, evaluate, not the total performance, but his sample of the performance, the performance as he perceives it. Finally, if this evaluation is to be recorded or fed back as information to the performer it must be converted into a message capable of being received by the performer, usually in a symbolic form such as language. A teacher's statement, such as "your posture was poor during your speech," must be recognized as information which has gone through so many reduction processes and channels since the original phenomenon, that it may provide useless or detrimental information to the performer and, consequently, should not be accepted unquestioningly if possible alternatives exist.

One alternative to other-evaluation is self-evaluation by the performer. Self-evaluation, either consciously or unconsciously, takes place whether other-evaluation is present or not. Self-evaluation of a psycho-motor performance may take place during the performance itself and afterwards through recall. Performer self-evaluation during the performance is based primarily on intrinsic feedback channels operating at the sub-conscious level, such as kinesthetic information as to the position of the limbs during a dance. Conscious attempts at performer self-evaluation during the performance could lead to an interruption of the performance. Performer self-evaluation upon completion of his performance relies upon recall for information. Recalled information can suffer from the fact that it is derived from a subjective point-of-view, that is, the performance as remembered by the performer, rather than as an effect upon the environment. Furthermore, the performance as a recollection must, because of the limited and selective capacity of the memory, contain much less

information than the original performance itself.

More powerful feedback to the performer than that coming the circuitious route through another evaluator and more complete and accurate than that coming from recall could be effected if the means were devised to record and play back the performance in an isomorphic medium, with objectively rather than subjectively transformed data. This study involved the use of a videotape recorder as a mirror in which the performer could see and hear his performance played back to him. The effect of this augmented feedback on self-evaluation and on subsequent performances and the difference between self-evaluation and other-evaluation of videotaped performances were explored in this study.

The use of the videotape recorder as an aid for self-evaluation has become popular in the last ten years with the development of inexpensive portable videotape recorders. This technique has been put to use in such fields as teacher training, psychotherapy, sports, and management and sales training, and has been acclaimed enthusiastically. Very little research, however, has been carried out into its effectiveness and the little evidence which does exist indicates that it has met with only limited success. There appears to be a serious lack of information at the public school level. With the videotape recorder becoming a more common instrument in the school, the results of this experiment may be useful to educators in helping them determine modes of utilization for this new medium. This thesis, then, sets out to explore the effectiveness of using the videotape recorder in one of its possible modes: as an aid for self-evaluation by elementary school children.

To summarize, then, this study was developed to meet several inter-related needs: the need for educational technology to apply itself to supporting new educational objectives based on giving the learner more control over his learning experiences; the need for developing the means for performer-evaluation of a performance as an element in learner-centred education; and the need to determine the effectiveness of using the videotape recorder as a medium to feed back information to a performer for self-evaluation.

CHAPTER III

REVIEW OF THE LITERATURE

The videotape recorder, particularly the relatively inexpensive half-inch recorder, is a recent innovation. The number of studies reporting its use in learning situations is consequently limited. A fairly extensive search of the literature of the 1960s produced a good number of monographs and journal articles which can be characterized as short, usually laudatory, descriptions of the use to which some educators have put their recorders. A much smaller number of articles report research carried out on the effectiveness of the videotape recorder in relation to self-evaluation and behaviour modification. Only one study was found dealing with the more specific problem studied in this thesis: the effectiveness of the videotape recorder at the elementary school level. As a result of this lack of information this review of the literature has had to go into fields beyond those encompassed by this thesis to obtain information which might be relevant to it. At the same time an attempt has been made to avoid those short, descriptive pieces of literature which provide little more than information about what people are doing elsewhere with videotape recorders. The review is divided into three areas: 1) the literature on feedback which provided the concepts and theory underlying this study; 2) the literature on self-evaluation not based on videotape systems; 3) the literature studying the influence and effectiveness of videotape feedback in selected fields.

Feedback

The literature on learning explained in terms of cybernetic models of behaviour is now quite extensive. Since this review has sections dealing with the more specific literature on self-evaluation and on the use of videotape recorders for feedback, this section will cover only those sources which influenced this study on a more general, theoretical level.

Peavy (1969) describes the cybernetic educational environment as one in which the learner is acting and transacting as a self-regulating organism rather than as a reacting one. In traditional educational processes, where learning is viewed primarily in terms of learner response, attention is focused on teaching, on the reactive processes in learning, and on the use of external guidance and motivational practices.

> Educational designs for the reacting learner stress reward, reinforcement, verbal persuasion, threat and punishment. Learning environments which are designed for the acting or transacting learner . . . . encourage choice, problem solution, and relevancy. In general they emphasize the self-regulating characteristics of a learner (Peavy, 1969, p. 15).

Defining the learner as a closed-loop system, one in which the output is returned to the system and consequently used to affect future outputs of the system, Merrill (1968) distinguishes between two types of feedback affecting the learning process: when the learner makes a response in his environment he receives at the same moment proprioceptive feedback indicating that the response has been made; when his response produces some change in the environment this change becomes a stimulus which is fed back to the learner as knowledge of results. Using these concepts learning is defined as

"the process of response modification as a result of proprioceptive feedback and knowledge of results . . . ." (Merrill, 1968, p. 6).

Annett (1969) takes much of what psychology has explained in the traditional stimulus-response theories and explains it meaningfully in cybernetic terms, in a psychology based on reaction to consequences rather than to causes. He distinguishes, as did Merrill, between intrinsic and extrinsic knowledge of results, noting that only the latter is easily subjected to experimental control. He also uses the term augmented feedback to describe the additional feedback loop that an experimenter may introduce into learning experiments, and notes that most augmented and extrinsic feedback is delayed temporarily from responses as opposed to intrinsic feedback which is generally immediate. Two other terms for the same concepts are action feedback, that feedback which arrives at the same time as, and can be used during, a response, and learning feedback, that knowledge of results " . . . which comes after the completion of the response such that information cannot be used to control the response being measured but can only be used by the subject for subsequent responses" (Annett, 1969, p. 29). Using Annett's terms to describe the processes involved in this study we see the person delivering a speech using intrinsic or action feedback during the performance and using the videotaped playback of that performance as extrinsic or learning feedback. The videotape recorder serves as an augmented source of information provided by the experimenter which cannot be used to alter the performance being recorded but which can be used to alter the subsequent performances. An important factor in both intrinsic and extrinsic feedback is the transformation rule,

the means of converting the measure of output into appropriate control action. In the case of the present study, the videotape recorder is no more than a device which displays electronic impulses on a tube. The perception, interpretation, and utilization of that information to alter future performances is governed by transformation rules used by the learner.

Smith and Smith (1966) in an extensive study of the cybernetic principles underlying learning regret the fact that the terms "feedback" and "knowledge of results" have become confused. They prefer to reserve the term "feedback" for the direct, dynamic sensory information that a performer receives during a performance. This is the proprioceptive feedback of Merrill (1968) and the intrinsic or action feedback of Annett (1969). They see "knowledge of results" as the static, terminal information received by a performer after completion of the performance. They accept, however, that the term "feedback" has come to be used for both types of information and that the similarities between the two are important.

> Either can serve to inform the individual about the accuracy of his movements. When a particular response provides no intrinsic feedback error signal, it often is critically important to give knowledge of accuracy in the form of an extrinsic signal at the end of the response. Verbal and symbolic learning often must be guided or defined in terms of static extrinsic knowledge until the individual has established intrinsic standards by means of which to monitor the learned responses (Smith and Smith, 1966, p. 208).

The term "feedback" is used in this thesis to refer to the information contained on videotape which is played back to the performer upon completion of his performance, and, as such, refers to Smith and Smith's "knowledge of results" or "knowledge feedback" rather than "dynamic sensory feedback."

Ammons (1956) surveyed the literature on knowledge of performance and produced several generalizations which are appropriate to this study.

1. The performer usually has hypotheses about what he is to do and how he is to do it, and these interact with knowledge of performance . . . .
. . . . . .
2. For all practical purposes, there is always some knowledge of his performance available to the human performer . . . .
. . . . . .
3. Knowledge of performance affects rate of learning and level reached by learning . . . .
. . . . . .
4. Knowledge of performance affects motivation . . . .
. . . . . .
5. The more specific the knowledge of performance, the more rapid the improvement and the higher the level of performance (Ammons, 1956, pp. 281-287).

Ammon's generalizations will be referred back to in some of the studies reviewed later in this section.

Self-Evaluation

Hagen and Thorndike (1960) in a review of the literature on evaluation found little evidence that student self-evaluation results in better learning. They reason that this failure is due to a number of factors:

> Clearly the student is unlikely to have much knowledge of techniques of appraisal, a clear understanding of the objectives of the school program, or a stable reference point for evaluating his own progress. His self-evaluation is likely to be based upon impressionistic opinion, unsupported by an accumulation of evidence. In addition, the social rewards for educational achievement introduce strong pressure for the student to distort his self-evaluation (Hagen and Thorndike, 1960, p. 485).

All these points are valid only if one accepts certain assumptions about the learning system in which that evaluation takes place. Basically, what Hagen and Thorndike assume, and what most of the

researchers who have studied self-evaluation assume, is that the learning system remains unaltered, remains primarily teacher-centred, after the introduction of self-evaluation. Given that assumption they proceed to show that self-evaluation does not work. They neglect to mention that little research exists to indicate that teacher-evaluation results in better learning and that teacher-evaluation suffers from most of the weaknesses attributed to self-evaluation. Most of the research judges the accuracy of self-evaluation using teacher-evaluation as a criterion. Even if one does not dispute the reliability of teacher-evaluation it is obvious that if one wishes to compare the accuracy of one measurement against another any conclusions as to which is more accurate will remain relative unless a third independent measurement exists as a criterion. The danger of bias entering into self-evaluation arises when evaluation systems are used for distributing rewards and punishments external to the learning tasks. The learner would be much less likely to distort his evaluation if its function were simply one of providing him with informative feedback about his performance rather than with the motivational feedback attached to it in learning systems which offer rewards for achievement. While Hagen and Thorndike explain the failure of self-evaluation by stating that "clearly the student is unlikely to have much knowledge of techniques of appraisal . . . .", as if this were some genetic trait, nowhere do they note that there has been little attempt made to develop the ability to evaluate in students.

Sawin (1969) notes that self-evaluation does take place in school learning. He points out that if learning were dependent

solely upon teacher evaluation, then students would learn only when in contact with the teacher as an evaluator.

> It is quite obvious that students do often learn without overt assistance from teachers. It follows, therefore, that students frequently do perform many evaluative functions. This means that the problem is not one of getting students to evaluate themselves, rather it is getting them to do it accurately and in such a way as to contribute most effectively to their learning progress (Sawin, 1969, p. 194).

Bloom, Hastings and Madaus (1971) make the distinction between summative and formative evaluation in learning for mastery. Summative evaluation takes place at the end of a learning unit and is used primarily for extra-system purposes, whereas formative evaluation is used as feedback to the learner and the instructing sub-system so as to appraise progress on the way to mastery. In a similar vein, Nuttin and Greenwald (1968) distinguish between success and reward in learning tasks: ". . . success is an intrinsic consequence of an action while reward is an extrinsic consequence. It follows that reward may satisfy some motive that is quite tangenital to successful performance . . . ." (Nuttin and Greenwald, 1968, p. 11). While Nuttin and Greenwald were not concerned primarily with the difference between learner- and teacher-evaluation, Nuttin's experiments on the varying effects of reward and success on attitudes towards the learning task offer some interesting evidence of the influence of other-evaluation. In one series of experiments two groups of seven and eight year olds were given an "intelligence" task in which they were able to see if they were successful or not. The tasks were administered in such a way that all the subjects were successful on half the tasks and unsuccessful on the other half. The subjects of one experimental group, in addition to having knowledge of results perceivable in the

task itself, were also informed verbally by the experimenter whether they were correct or not. Despite the fact that subjects in both groups had the same number of successes and failures a significant majority in the first group recalled a higher level of failure than had really occurred, while a majority of the second group recalled a higher level of success than had in reality occurred. Interpreting Nuttin's results, it would appear that teacher feedback, even though accurate, can result in a greater feeling of failure on the part of the student being evaluated, while independent self-evaluation can result in a greater feeling of success.

Personal characteristics of the students may also affect the evaluation process. Russel (1953) reported one study which found fifth- and eighth-grade boys evaluating themselves lower than girls, with both of these ratings significantly higher than teacher ratings. In another study, also mentioned by Russel, fourth-, fifth-, and sixth-graders rated themselves higher than did their teachers in reading, spelling, and handwriting, with a positive correlation between the two evaluations being shown. Pupils of high intelligence underestimated and those of low overestimated their performance in this study.

Videotaped Feedback

A search of the educational indexes, the Educational Resources Information Center (ERIC)collection, and a computer search of the Dissertation Abstracts from 1960 to 1970 located but one study (Hurley, 1971) of the effect of videotaped playback on pupil evaluation at the elementary school level. A fair amount of literature exists reporting on the usage of videotape recorders at other levels since the early 1960s, the date around which inexpensive videotape recorders

became available. Most of this literature is non-research based, being mainly reports on how videotape recorders were used and how well they were received. The few controlled studies carried out on the effectiveness of the videotape recorder as a feedback mechanism have, on the whole, not proved the superiority of this medium in comparison to other forms of feedback. Since little literature was found relating directly to the problem under study here the wider ranging literature reported below was included so as to provide some background to this study. This section will examine some of the studies on the effects of videotapes carried out in psychotherapy, business training, sports, teacher education and public speaking after first reviewing a study using film as the basis for self-confrontation.

An important factor to consider in using the videotape recorder as a feedback mechanism is that the information flowing back to the performer is not emotion-free information but is a presentation of himself, a self-confrontation, that is novel and rare in its occurrence. The videotape recorder is a mirror for the performer. One of the major studies on the effect of such audio and visual self-confrontation was carried out by Nielson (1962) using film. The use of film resulted in a one-week delay between recording and play-back, a delay not necessary with videotape. Nielsen was interested in personality, in studies of self-awareness, rather than in self-evaluation of performance. He found that in viewing films of themselves embroiled in argumentative dialogue, college student subjects ". . . were in a state of intense, sometimes painful self-awareness, different from any normal, daily self-awareness" (Nielsen, 1962, p. 27). "Some of the subjects were curious as to the nature of their

inner self; others were frightened by the intense self-awareness and found it quite painful. The confrontation with the self-image left none of them neutral or untouched" (Nielsen, 1962, p. 41). Nielsen measured the amount of self-attention, the amount of time subjects fixated upon themselves on the screen, and found a wide range (30-80 per cent) among the subjects. A positive correlation was found between measured narcissism and self-attention. The subjects re-confronted themselves one and a half years later with interesting results: their measured self-attention increased but retained the same rank order; they became more detached, more objective in their evaluation of themselves and of the opponent in the film; those who had evaluated themselves lowly in initial self-confrontation now evaluated themselves higher, while those who had high initial evaluations of themselves were more negative on re-confrontation; evaluation of their opponent shifted from a view of him as a type to one of him as an individual, while self-evaluation showed an opposite trend with subjects on re-confrontation seeing themselves more as types, as examples of being at an age-stage, rather than as individuals. Nielsen made no attempt to measure behavioural change but some subjects volunteered that they had changed because of their experiences in self-confrontation. They reported changes at the level of ridding themselves of minor mannerisms.

Ryan (1968, 1969) exploring the differences between videotape recordings, film and broadcast television notes that film, because of the delay in processing, contains information useful in the future, often to someone else, while videotape

> . . . gives you access to the data of your immediate environs as well as access to yourself in a cybernetic way. There is a closed circuit: a feedback loop with oneself available through

> videotape. Your output or behaviour is input for the video-recorder to process. What it puts out on its monitor is input for you to process into modified behavior and so on. Videotape grants access to one's own body language to offer a unique opportunity for self-processing (Ryan, 1969, p. 30).

Nor, he notes elsewhere, is videotape like broadcast television: "VT is not TV. If anything, it's TV flipped into itself. Television, as the root of the word implies, has to do with transmitting information over distance. Videotape has to do with infolding information" (Ryan, 1968, p. 38).

The possibility of using video feedback as a therapeutic tool in psychiatry has led to its being adopted as a means of giving patients a more objective view of themselves. Alkire (1969) in reviewing the literature on the use of videotaped playback in education and mental health distinguishes between the feedback we receive daily in our routine interactions and the feedback we receive from a videotaped replay of our own performances:

> . . . the video monitor is a source of information whose motivations can only be "pure". There is no basis for doubting what it reveals about ourselves on the grounds that it is biased, or has something to gain by distorting facts . . . . We may be genuinely surprised then at how clearly we can see matters when we <u>cannot</u> resort to questioning the judgment or motives of other persons as a source of negative information about ourselves (Alkire, 1969, p. 189).

Alger and Hogan (1969) report that valuable residual effects were produced as a result of the use of the videotape recorder in therapy with a feeling of mutuality being developed between the patient and his therapist as the patient became a co-diagnostician. Boyd and Disney (1967) found that after only one exposure to video feedback, mental patients' concepts of self, ideal self and public self became less discrepant and pathological compared to a control

group and remained lower even two weeks later. Another study found that the counselling client was able to view himself objectively, as " . . . being well known to him, and yet not quite he" (Kagan, Krathwohl and Miller, 1963, p. 239). Stoller (1967) reports research indicating that after self-confrontation via videotape a person will first change his overt behaviour before incorporating the change into his self-concept.

Another field in which the videotape recorder is being used as a feedback tool is in business and sales training. One of the few studies based on research carried out in this field was by Stroth (1969) who found no significant difference between two groups of industrial salesmen, one of which used videotape and the other audio-tape as a feedback mechanism. Stroth cautions that " . . . videotape users have been so impressed by the extensive student interest generated by self-confrontation that they have concentrated on determining the degree of interest rather than the degree of learning" (Stroth, 1969, p. 19). As noted earlier by Ammons (1956), the providing of knowledge of performance can in itself be motivating and while this may not be an undesirable effect, the satisfaction generated by using the videotape recorder should not be confused with its effect in producing changes in performance when evaluating the effectiveness of video feedback. Although Stroth found audio-taping to be as effective as video-taping it should be noted that he was measuring changes in the conversational behaviour of salesmen trainees and used as a criterion measure the amount of time spent in speaking. Since the only element of performance being measured was an audio component, his findings are not surprising. Stroth does not test the videotape

recorder in the field it can best exploit, in its effect on the visual components of behaviour. An interesting finding of Stroth's was that audio-taped trainees improved their performance consistently over three trials whereas the videotaped group's performance declined on the second trial but caught up to the audio-taped group on the third trial. Stroth expressed regret that further trials beyond the third were not possible in his research.

Studies which exploit the video component in videotaped feedback may be found in the field of sports training. DeBacy (1970) studied the accuracy of self-assessment of a golf swing by students in a beginning golf class. The assumption that subjects can perceive accurately underlies all the research in this field. Using the ratings of judges as a measure of actual skill, DeBacy found that the beginning golf student made inaccurate self-assessments in the direction of over-assessment, that the viewing of a model performance did not increase accuracy, but that the viewing of a model performance and one's own performance on videotape produced a reduction in over-assessment such that the self-ratings after video replay were not significantly different from the judges' ratings. Another study using judges' ratings and self-ratings of bowling skills, (Robb and Teeple, 1969) found that students did not alter their ratings after viewing themselves on videotape but that the judge's ratings of the video-taped performance was significantly different on two out of seven measurements from the same judge's ratings of the live performance with the result that discrepancies between the self-ratings and the judge's ratings which existed prior to the video replay disappeared after replay. This piece of research involved only thirteen subjects, one judge, and

no control group, but it does indicate that evaluation even by a person other than the person being evaluated is not a static event but a process subject to change.

Perhaps the field which has made the greatest use of videotaped playback as an aid for self-evaluation and performance modification is the field of teacher-training. Videotaped playback has become an integral part of a teacher-training technique called microteaching. Developed by Cooper and Allen (1970), microteaching involves the teaching of a short lesson emphasizing specific teaching skills and following this by immediate feedback which may come from video- or audio-tape recordings, supervisors, pupils, colleagues and self-evaluation by the student-teachers. The lesson is often repeated immediately so as to incorporate any changes resulting from the feedback.

Shively (1970) set up a microteaching situation in which the feedback mode was varied and found that audiotaped feedback and pupil-evaluation feedback were superior to videotaped feedback in producing change. Videotaped feedback, however, was preferred by the participants. A multi-skill inventory measuring instrument, the Stanford Teacher Competence Appraisal Guide, completed by the pupils being taught the lesson served as the criterion measure. It was used as the feedback mechanism in one of the experimental groups as well. This, coupled with the fact that the instrument has a strong verbal orientation, might explain the superiority of the pupil- and audiotape-feedback modes. Furthermore, only one re-teach session was held, whereas, as we have seen, Stroth (1969) found that videotaped feedback produced a detrimental effect on the second performance when compared to audiotapes, but became more effective by the third trial.

Ward (1970) also varied the mode of feedback and found that, if the re-teach session followed immediately after feedback, no change in teaching performance, as measured by the number of questions asked by a teacher, resulted from the use of either videotape alone, videotape and model performances, audiotape alone or reflective self-evaluation. When a one-week delay was placed between the feedback and the re-teach session, the videotape and audiotape feedback modes produced greater changes in performance. Once more, however, the criterion measure was verbal and thus only exploited the audio channel of the videotape recorder.

Jensen (1968) assigned 65 elementary school teachers to five treatment groups over a two-year period and found that the groups using videotaped-based self-evaluation exhibited moderate changes in teaching methods and attitudes compared to those who used unaided self-evaluation. He found them willing to recognize negative information of their performance without developing defensive or defeatist attitudes.

Solomen and McDonald (1968) in reviewing the literature on self-confrontation distinguish between two types of studies. Some studies showed changes in behaviour resulting from self-confrontation while others reported rejection or evasion of the feedback messages. They found two conditions were present in the former group of studies: 1) the receivers knew what behaviour was expected, and 2) they adopted these expectations and were willing to adapt their behaviour to coincide with these expectations. The authors claim that only when these two conditions were met could the information properly be called feedback. When no standards were set, however, other variables

determined by the receivers' self-perceptions and predispositions came into play. An experiment in teacher self-evaluation was set up in which neither of the conditions were met, that is, in which the participants were given no standards with which to judge themselves other than the standards which they might have brought into the experimental situation. It was hypothesized that given this absence of standards, attitudinal changes and perception of cues would be determined by the teachers' satisfaction with their performance. The subjects were interviewed and given an attitude questionnaire prior to the videotaped playback of a fifty minute teaching performance. The procedure was repeated after playback. As hypothesized, low-satisfaction teachers devalued concepts in the institutional domain after self-confrontation while those with high-satisfaction increased their evaluation of their professional selves. The majority of cues attended to were related to physique with the low-satisfaction group eliciting more responses in this class than the high group. The high-satisfaction group attended more to teaching behaviour cues. The authors conclude that:

> . . . when no model of "good teaching" is presented, no guidance is given, and no new and common standards are adopted, reactions to self-viewing on videotape are determined largely by the viewer's predispositions. That is, his contentment with his own performance determines what will be noticed on the screen, how it will be evaluated and to what attitudinal change it will lead (Solomen and McDonald, 1968, p. 15).

These findings coincide with Ammons' (1956) generalizations concerning knowledge of performance and can be related to Annett's (1969) concept of the transformation rule. The performers had some knowledge of their performance prior to the playback of their videotapes and this knowledge, measured in terms of satisfaction with their

performance, given the absence of the other standards, influenced the selection of cues and criteria used in evaluating the video playback. Solomen and McDonald did not try to repeat their experiment with a group using standards of performance nor did they attempt multiple sessions with the same teachers to see if attitudes and behaviour would change on subsequent performances as the cosmetic effect of the first wore off and as their own standards came to the fore. Their work is important, however, in pointing out that the mere replay of a videotaped presentation is not what is meant by feedback, but rather that feedback is that information taking an evaluative form, that information processed in terms of its relation to a standard.

The last area to be reviewed here in which videotaped replay has been used for self-evaluation is speech delivery. In this field, as in the others mentioned so far, no literature was found dealing with elementary school students. As in the other fields the majority of the literature consisted of short laudatory journal articles of a descriptive rather than a research nature.

One of the first published accounts of the use of videotape recordings and the first to use the term "mirror television" was by Henderson (1964) who explains that the term indicates the role that television plays in the type of learning it generates, one of " . . . simply reflecting or mirroring the subject's performance . . . where the TV instrument is made a tool for the learner, as contrasted with when it serves primarily the aims of the instructor in the act of teaching and presenting subject matter" (Henderson, 1964, p. 53). Nelson (1968) found that using a videotape recorder in college speech courses was well received by students and resulted in

more speech preparation. Becker, Bowers and Gronbeck (1968) used a studio set-up rather than a classroom and super-imposed the instructor's comments and visuals on to the videotapes as reinforcement and guides. They reported that students liked this approach and that it seemed effective. None of these studies, however, reported any experimental data to back up their claims.

Hirschfeld (1968) studied the difference between self-, peers-, and judges-evaluations and found that peers judged performances higher than did judges, with self-evaluations based on videotaped playback falling between the two. Although no tests of significance were undertaken, it was reported that the three types of evaluations were all close to each other. The only exceptions were that in self-evaluations the speakers were more aware of stage-fright and speech accents. Dieker, Crane and Brown (1968) found that students receiving videotaped playback evaluated themselves closer to the instructor's ratings than did a control group with no video replay. A study by Frandsen, Larson and Knapp (1968) of the effect of mixing videotaped playback and intructor's comments on bringing about correspondence between the speakers' self-evaluations and the instructor's evaluations found that the greatest correspondence occurred when videotaped playback preceded the instructor's comments with less correspondence when these preceded the replay or were given simultaneously with the playback.

Mulac (1968) undertook a study to compare the relative effectiveness of videotape, audiotape and traditional (class and instructor) feedback to test a general hypothesis that "the greater the completeness and accuracy of student speech performance feedback

the greater the degree of speech skill a student will later exhibit" (Mulac, 1968, p. ii). College students in a freshman speech course, stratified by age, sex, speech background and instructor, and randomly assigned to three feedback groups, were judged by nine judges on a multi-factor performance scale. The results indicated that the videotaped groups showed significantly ($p<.01$) greater overall speech skill than the audiotaped and traditional groups with a 40 per cent greater gain in the criteria of bodily action, personality, language, and voice. Describing feedback completeness or accuracy in terms of the number of sensory stimuli inherently transmittable by the recording-replay device, Mulac explained that the alternative to videotaped playback

> . . . requires reliance upon the verbal descriptions, evaluations, and perhaps mimicking gestures and vocal inflections of a classmate or teacher. In such a situation the student often fails to fully understand the nature of the criticism or to fully accept its validity. With electronic replay, on the other hand, he can view the performance for himself. Even if his reaction to his own performance is not completely "objective", the precision and completeness of the feedback should lead to a more accurate picture and greater motivation than traditional criticism methods (Mulac, 1968, p. 13).

Only one study of the use of videotaped feedback with elementary school children was found. Hurley (1971), in a design closely approximating the one used in the present study, assigned a group of forty subjects to four feedback treatments: 1) videotaped playback of performance, 2) videotaped playback with checklist-guided self-evaluation, 3) checklist-guided self-evaluation alone, and 4) exposure to neither of the manipulated variables. The tasks consisted of non-verbal psychomotor operations (layout, holding, cutting, and assembling) performed twice with the experimental treatments being

applied between repetitions. The second performance was videotaped and rated by three judges. The results indicated that those subjects receiving video playback while rating their own performance on the self-evaluative checklist performed on a higher level than the control group in three of the four tasks. Subjects using the checklist alone performed higher than the control group on two tasks while those having only videotaped feedback performed higher on one task. Unlike the present study, Hurley's did not compare first trial to second trial performance for any of the groups, nor trace performances over several trials, nor compare self-evaluation to the judges' evaluations but it did partially confirm the hypothesis common to both studies that more complete feedback would result in higher performance.

In summary, then, the literature reviewed here provided this study first with a vocabulary and conceptual framework. The subjects in this study were defined as self-regulating learners transacting with their environment (Peavy, 1969). While all subjects were receiving proprioceptive or intrinsic feedback, some were exposed to augmented feedback or knowledge or results (Merrill, 1968; Annett, 1969) via the videotaped playback of their performances. The studies on evaluation indicated that a distinction between summative and formative evaluation (Bloom, Hastings, and Madaus, 1971) and between success and reward (Nuttin and Greenwald, 1968) was important. The research in this area also indicated that the results and consequences of self-evaluation could differ significantly from those of teacher-evaluation. Studies of the effect of providing the learner with a videotaped replay of his performance indicated that this resulted in a more objective self-evaluation (Alger and Hogan, 1969; Alkire, 1969; Debacy, 1970;

Jensen, 1968), or an evaluation closer to the instructor's (Dieker, Crane, and Brown, 1968; Frandsen, Larson, and Knapp, 1968), provided that the learners were sufficiently cued as to what behaviour was expected of them (Solomen and McDonald, 1968). Studies which indicated that the videotape recorder was not significantly superior to traditional and audio sources of feedback (Stroth, 1969; Shively, 1970) in altering student performances were found to emphasize the audio rather than video components of behaviour. A study (Mulac, 1968) which considered non-verbal behaviour found the videotape to be superior to the audiotape recorder and to traditional modes of feedback while a study (Hurley, 1971) approximating in design and subjects the present study, confirmed the hypothesis that videotaped playback accompanied by a self-evaluative checklist resulted in a higher level of performance.

EXPERIMENTAL DESIGN

CHAPTER IV

HYPOTHESES AND QUESTIONS

The theoretical basis upon which this research lies is a cybernetic one which sees the behaviour of an organism as a reaction to information fed back to it from its internal and external environment. Learning is seen as a process which alters future performances of an organism based on an evaluative reaction to the information received from past performances. Underlying this theory is the assumption that the more complete and accurate the feedback received, the more likely learning is to occur.

The general hypothesis tested here is that, in addition to the proprioceptive and recalled information present in a performance, the application of two independent variables, 1) video-tape replay of a performance, and 2) a self-administered evaluative checklist, presented singly and in combination immediately following a speech delivery, constitute an example of more complete and accurate feedback which will manifest itself through an improvement in subsequent speech delivery performance as measured by two dependent variables, 1) judges' evaluations and 2) speaker self-evaluations.

The specific hypotheses to be tested in this thesis are listed below.

H-1: Subjects in groups CHKL, VTR-CHKL, VTR and CNTL will exhibit statistically significant improvements in their speech delivery performances over five trials as measured by three judges viewing

videotapes of these performances. (It is expected that improvement will occur regardless of experimental treatment as a result of the non-manipulated sources of feedback present for all groups repeating the same task five times.)

H-2: Subjects receiving augmented feedback of their speech delivery performances will exhibit significantly greater improvements in performance as measured in the interaction effect between type of experimental treatment and speech trial over five trials, as measured by three judges viewing videotapes of these performances. The differential improvements in measured performance will be exhibited in the following relationship:

VTR-CHKL > VTR > CHKL > CNTL

H-3: Subjects using the self-evaluative checklist will exhibit statistically significant improvements in speech delivery performance over five trials when measured by their own evaluations. (This is analogous to H-1 but refers to self-evaluation.)

H-4: Subjects receiving videotaped playback of their speech delivery performances will self-evaluate their own performances differently from those not having such replay such that VTR-CHKL $\neq$ CHKL. (No direction is hypothesized at this time although the results will indicate the relative size of these scores if the hypothesized inequality is confirmed.)

H-5: Subjects receiving augmented feedback in the form of videotaped playbacks of their speech delivery performances will produce self-evaluations closer to the judges' evaluations of their performance than those not receiving such feedback, such that the scores generated by subtracting the judges' evaluations from self-evaluations for groups

VTR-CHKL will be smaller than those generated for group CHKL.

H-6: Subjects receiving augmented feedback will perceive more in their speech performances, as measured by a content analysis of their oral self-evaluations, than those not receiving such feedback according to the following relationship:

VTR-CHKL >VTR >CHKL >CNTL

In addition to testing the above hypotheses in a statistical mode, the following questions will be answered in a descriptive form.

Q-1: Do the subjects identify with the video and audio images of themselves during playback?

Q-2: How do the subjects react during playback of a videotaped presentation of themselves?

CHAPTER V

DESIGN OF STUDY

Since little published research was found on the use of videotaped feedback with elementary school children, the design for this study had to grow, then, more out of experience than from previous research. Pilot studies were undertaken to decide which psycho-motor skill was to be taped and evaluated, to try out possible techniques, and to develop the measuring instruments and procedures to be followed in the experiment. Given this origin, it would seem useful to describe here, in short form, some of these pre-experimental experiences which resulted in the particular experimental design to be described later.

Pilot Studies

The investigator visited Crestview Elementary School twice a week from October 1969 to June 1970 and had many opportunities to meet with the pupils and their teachers and to use the videotape equipment under a variety of situations. The process of recording and playing back pupil performances was used in physical education, special classes in psycho-motor coordination, language arts, French, kindergarten play activities and the film club. The investigator found the highest teacher demand for his services in the language arts classes for drama, public speaking, debates, and "show-and-tell" at all grade levels. In seeking a psycho-motor activity requiring fairly complex and identifiable skills with overt behaviour which would exploit both the audio

and video components of the videotape system while being relevant to the school curriculum, public speaking was selected for experimental treatment.

The first series of pilot studies in public speaking took place with senior pupils in the classroom. Speeches were delivered to the class by several pupils in succession. About halfway through the period the tape was played back. Using this technique allowed only a few pupils in any one session to be exposed to the process of videotaped feedback, and none the opportunity to repeat his performance following feedback. The investigator found that under these conditions he had little control over or access to, self-evaluations and reactions to self confrontation.

In the second series of pilot studies a few of the senior pupils, selected by their teachers, were recorded delivering speeches individually and in groups of two or three outside the classroom environment. An important element of public speaking, the audience, was lost but the experimenter's access to the individual speaker's reactions to self-confrontation was increased. The most obvious difference between this series and the earlier one was in the speakers' overt reactions during playback. In the classroom initial self-confrontation was accompanied by much giggling, heavy facial blushes, embarassed covering of the face, and whispered comments to friends on the part of the speaker; laughter, occasional ridicule, verbal and non-verbal compliments and consolation, occasional boredom and disinterest were exhibited by the other viewers. In the second series, when the speaker was alone with the investigator, overt reactions were much less intense in terms of gross body movements and verbalizations. Intense facial concentration,

immobile, stiff bodies, and reluctant and sparse verbalizations were the rule. When one or two other pupils were present the reactions were a mix between the two types described above, but were closer to the second, those occuring when the pupil was alone.

This second series allowed access to the speakers' verbal reactions to their speeches. It was found that they had little to say in evaluation. This reluctance to speak seemed inconsistent with the intense reactions exhibited. When the investigator attempted to elicit evaluations through questions relating to the speeches, this resulted in an increase in verbalized evaluations, but it was evident that the investigator's own reactions to the individual speech could contaminate the speaker's self-evaluations.

A third series of pilot studies was carried out under the same circumstances as the second. This last series was devoted to developing the procedures to be used in the experimental situation. The second series had pointed out the need for some tool which the subjects could use in evaluating their own performance. A speech check-list was developed during this series which gave the subjects criteria upon which to judge their speeches while still allowing the evaluation to come from the subject with a minimum of other-person influence. This checklist was revised during this series so as to overcome ambiguities due to difficult or inadequate wording and became the Speech Ruler used in the experiment.

Data Collection Procedures: Subjects

The subjects participating in the experiment were chosen at random from the total intermediate level population at Crestview Elementary School as described in Chapter VI. The random selection and

assignment to treatment groups was carried out without the subjects' knowledge. They were not told that an experiment was involved but merely that the activities would be part of their language arts programme.

A week before the experimental sessions began the investigator visited the intermediate level classes to explain what would be involved. The students were told that they would be called from mid-May to the end of the school year to deliver a speech to be recorded on videotape. The reason given to them for this activity was to offer them the opportunity of seeing themselves deliver a speech, an opportunity which only the senior students had heretofore enjoyed during the year. They were further told that in order to allow each pupil the maximum amount of time on videotape, they would be selected in groups of two with one week's notice to allow them time to prepare a speech. Five topics for speeches, selected in consultation with the teachers in terms of interest and open-endedness, were recommended. Each pupil was then given a mimeographed summary of these general instructions. Other copies were placed on the bulletin boards in the intermediate area. A copy of this handout may be found in Appendix A.

Each week a list of names of the subjects to be taped in the following week was posted on the bulletin boards. Further notice was given to each subject one day before the taping session in the form of a note with the subject's name and the time at which he should appear for the taping.

The taping sessions took place on the stage in the gymnasium with the curtains closed blocking out the sight and most of the sound from the gymnasium. The equipment set-up included a videotape recorder and monitor on a mobile table next to the camera. The

microphone was set up twenty feet away from the table. The subjects were instructed to stand next to the microphone, a piece of tape marking the spot on the floor. A single floodlight located on the floor at a 30 degree angle to the camera-speaker line was aimed at the subject. At a right angle to the camera-speaker line and about ten feet from the monitor table were two desks and chairs where the subjects and the experimenter would sit during playback. The experimenter's desk was set next to the subject's desk but at a slight angle. This allowed the experimenter to appear to be looking at the monitor while actually observing the subject's reactions to the video playback. A portable cassette audio tape recorder was placed inside the subject's desk with a microphone extending out the ink-well hole. This was used for recording oral evaluations and verbal reactions to initial self confrontation. A sketch of the physical layout can be found in Appendix G.

When the subjects arrived for the taping of their speeches they were not told that the procedures to be followed were part of an experiment, nor that other subjects had been exposed to differing treatments. They were told that each subject was to deliver his speech five times. The purpose for this, it was explained, was to capture on tape the pupil at his best. It was explained that one delivery alone might not be a good example of performance and that several "takes" were necessary in television. The experimenter explained, however, that the last speech would not necessarily be the best but that other subjects had experienced their best speech on anywhere from the first to the fifth take. The purpose of these explanations was first to give the subjects a reason for delivering their speeches five times and secondly to avoid placing a "set" in their minds that delivery

would necessarily improve with consecutive trials.

The subjects were told that they would be asked to evaluate their own performances. It was explained that this evaluation might help them locate the strong and weak points in their performances and subsequently help them in the next performance. The experimenter emphasized that he would in no way try to judge their speeches because he wanted them to tell him what they though of their own speeches. The experimenter developed a routine of making a small smile and saying "good" after each speech as it was felt that no expression at all might be interpreted as dissatisfaction. Furthermore, the pilot studies had shown that many subjects would be nervous and apprehensive and that some form of consolation was necessary. Because two subjects were present, each was instructed not to tell the other what he thought of his speech. They were told that the time for sharing experiences would be after the sessions.

Following these general directions and a description of the television equipment for those who were interested, the experimenter collected the background data required to identify each subject. This was collected using the Data and Observations No. 1 form, a copy of which can be found in Appendix B. Some information was collected later from the school's files.

Up to this point all subjects had received identical introductions and instructions. They had, without their knowledge, been assigned to different treatment groups. After the general introduction and data collection, the subjects received instructions and took part in procedures specific to the treatment group they were assigned to. An outline of the four different sequences which subjects were

exposed to follows:

Treatment: VTR-CHKL

1. The experimenter (E) instructed the subjects (S) on general procedures for this group.
2. E instructed Ss on procedures for speech delivery for videotaping (e.g. where to stand, etc.).
3. Ss delivered speech No. 1.
4. E introduced and explained the Speech Ruler to Ss (see Appendix C).
5. Playback of speech No. 1.
6. E questioned Ss on initial self-confrontation and Ss evaluated speech orally (question 21 on Data and Observations No. 2).
7. Ss evaluated speech No. 1 using the Speech Ruler.
8. Ss delivered speech No. 2. Playback of speech No. 2.
9. Ss evaluated speech No. 2 orally.
10. Ss evaluated speech No. 2 using the Speech Ruler.
11. Steps 8, 9, 10 repeated for speeches Nos. 3, 4, 5.

Treatment: VTR

1. E instructed Ss on general procedure for this group.
2. E instructed Ss on procedures for speech delivery for videotaping (e.g. where to stand, etc.).
3. Ss delivered speech No. 1.
4. Playback of speech No. 1.
5. E questioned Ss on initial self-confrontation and Ss evaluated speech orally (question 21 on Data and Observations No.2).
6. Ss delivered speech No. 2.
7. Playback of speech No. 2.
8. Ss evaluated speech No. 2 orally.
9. Steps 6, 7, 8 repeated for speeches Nos. 3, 4, 5.

Treatment: CHKL

1. E instructed Ss on general procedures for this group.
2. E instructed Ss on procedures for speech delivery for videotaping (e.g. where to stand, etc.).
3. Ss delivered speech No. 1.
4. E instructed and explained the Speech Ruler to Ss (see Appendix C).
5. Ss evaluated speech No. 1 orally.
6. Ss evaluated speech No. 1 using the Speech Ruler.
7. Ss delivered speech No. 2.
8. Ss evaluated speech No. 2 orally.
9. Ss evaluated speech No. 2 using the Speech Ruler.
10. Steps 7, 8, 9, repeated for speeches Nos. 3, 4, 5.
11. Playback of selected speech or speeches; E questiones Ss on initial self-confrontation (question 21 on Data and Observations No. 2).

Treatment: CNTL

1. E instructed Ss on general procedures for this group.
2. E instructed Ss on procedures for speech delivery for videotaping (e.g. where to stand, etc.).
3. Ss delivered speech No. 1.
4. Ss evaluated orally speech No. 1.
5. Repeat steps 3 and 4 for speeches Nos. 2, 3, 4, 5.
6. Playback of selected speech or speeches; E questioned Ss on initial self-confrontation (question 21 of Data and Observations No. 2).

In summary, then, procedures in each treatment group varied according to the absence or presence of two independent variables: 1) videotaped playback after each performance and 2) use of the Speech Ruler with which to evaluate the performance. Subjects in all

four treatment groups received identical introductions and initial instruction; received identical instructions with reference to delivering a speech before a television unit (step No. 2 in each outline above); delivered the first speech (step No. 3) without previous exposure to videotaped playback, or the Speech Ruler, or instruction in speech delivery within the experimental setting; delivered five speeches and evaluated each speech orally. Furthermore, all subjects were exposed to self-confrontation via videotaped replay of their speaking performance, those in VTR-CHKL and VTR during the experiment and those in CHKL and CNTL after the five speeches were completed.

Data Collection Procedures - Judges

Each videotaped speech by each subject was evaluated by three judges in the summer following the taping. Two of the judges were ex-elementary school teachers. One was completing a Masters degree in English, the other was involved in community affairs and the free school movement. The third was a student in the Educational Technology programme at Sir George Williams University with extensive media-oriented experience with school children. None of the judges knew any of the subjects.

The judging sessions took place over a one-week period in a classroom at Sir George Williams. The videotape recorder and monitor were set up in front of the three judges' desks. Each speech was given a code number relating to, but different from, the code given to each subject, since it was possible to determine the treatment groups and speech number using the latter code.

The judges were instructed to consider each speech as a unique manifestation and not to try to relate it to other speeches by the same subject or to other subjects' speeches. The speeches were presented in a modified random order with those on any one tape being played back in a random order such that it would be difficult for the judges to determine which of the five speeches a subject was delivering. A completely random presentation of all the speeches on all the tapes would have been too time consuming as this would have involved the changing of tapes between speeches, there being twelve one-hour and half-hour tapes with twenty-four to fifty-five speeches per tape.

At the beginning of the judging sessions, the judges were given a general description of the purpose and procedures used in the experiment and the role they were to play in it. The judges were introduced to the checklist they were to use in evaluating the speeches. This was a modified version of the Speech Ruler and can be found in Appendix E. In instructing the judges in the use of their checklist, the investigator distributed copies of the Speech Ruler and repeated verbatim the instructions given to the subjects in its use. These can be found in Appendix C. The judges were warned against the dangers of halo effects between speeches by the same subject and between speeches by different subjects.

The judges were asked to share their evaluations on the first dozen or so evaluations so as to determine if all were using the same concepts in their evaluations. They were not discouraged from sharing their evaluations during the rest of the judging sessions. At the same time, however, the investigator did not discourage any judge from deviating from the others if such evaluations were felt

to be true. In this manner it was hoped that inter-rater reliability would be raised by assuring to a certain extent that the judges were in agreement on the meaning of the checklist's criteria while at the same time recognizing that the evaluation of a complex psycho-motor skill such as speech delivery involved a certain degree of subjectivity. A measure of inter-rater correlation was carried out to determine the degree to which the judges did agree with each other. A random sample of thirty speeches were extracted from the total 350 (70 subjects x 5 speeches each) speeches evaluated by the judges. The judges' evaluations were subjected to analyses to produce the following product-moment coefficients of correlation and multiple correlation:

$J_1 : J_2 = .81$ $\qquad$ $J_1 : J_2, J_3 = .82$

$J_1 : J_3 = .71$ $\qquad$ $J_2 : J_1, J_3 = .84$

$J_2 : J_3 = .74$ $\qquad$ $J_3 : J_1, J_2 = .76$

CHAPTER VI

DESCRIPTION OF SAMPLE

The subjects in this experiment were pupils at the intermediate level in Crestview Elementary School in Chomedy, Laval, a suburb of Montreal. Crestview is a public school under the jurisdiction of the Protestant School Board of Greater St. Martin. It serves a predominantly Jewish lower-middle and middle-middle class population living in small, single dwelling houses and duplexes. The school is involved in a non-graded and team-teaching approach to education with the pupils grouped in kindergarten, primary, intermediate and senior levels rather than grades. The intermediate level consists of pupils ranging in age from eight to eleven years and corresponds to grades three and four in graded schools. The 180 pupils and six teachers at the intermediate level work in an open area and four adjoining classrooms. Reading and mathematics instruction is individualized with each pupil progressing at his own rate.

The intermediate level was chosen for this study as it included a large population of pupils in the school, 1) who had not been exposed to the videotape recorder in the pilot studies carried out prior to the collection of the experimental data, 2) who had had varying amounts of formal training in speech delivery, 3) who had been less exposed to formal summative evaluation than older pupils and who would hopefully be less apt to think of evaluation in terms of "marks that

count."

Selection of subjects for the experiment was achieved by assigning each pupil a three-digit number from 001 to 180 and selecting from a table of random numbers a sample of 100 subjects. The three-digit number became a part of each subject's identifying code number. Each of the selected numbers was placed on a file card. These cards were shuffled and distributed into four packs of twenty-five. The letters A, B, C and D, corresponding to the four treatment groups, CHKL, VTR-CHKL, VTR and CNTL respectively, were each assigned to one of the four packs and became part of the subject's code number. Thus, subject B166 was a subject in the treatment group receiving both videotaped playback and the use of the self-evaluative check-list, that is group VTR-CHKL.

Although it was originally planned to administer the experimental treatments to four groups of twenty-five subjects each, a combination of circumstances resulted in an uneven distribution of subjects per treatment group. By the end of the data collection period the four groups contained 21, 20, 15, and 14 subjects each. The reasons for this uneven distribution include an earlier than forseen halting of the experiment due to the interference of many extra-curricular activities at the end of the school year, the elimination of some subjects who did not appear for the sessions or who could not complete the session in one sitting, and the accidental erasure of some tapes.

The composition of the four treatment groups in terms of their numbers and the distribution of certain characteristics for which data was collected may be found in Table 2.

TABLE 2.--Distribution of subjects in four treatment groups

Treatment	CHKL	VTR-CHKL	VTR	CNTL	TOTAL
Number of Ss	20	21	14	15	70
Female	10	11	9	7	37
Male	10	10	5	8	33
Previous Speech Training - Yes	9	11	5	7	32
Previous Speech Training - No	11	10	9	8	38
IQ Mean	96.05	102.00	95.57	102.60	99.05
IQ Median	97.00	100.00	98.50	98.50	99.50

Subjects had been assigned randomly to the four treatment groups. No attempt had been made to match the groups according to characteristics. The resultant uneven distribution of males and females within each group and of subjects with and without previous speech training within each group was subjected to chi-square analyses. It was found that in no group did this uneven distribution exceed expected frequencies (50 per cent on each trait) significantly. The mean I.Q.s of each group were subjected to an analysis of variance, reported in Table 3, which indicated no significant variation among the four groups.

TABLE 3.--Analysis of variance of mean I.Q.s of four treatment groups

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	F	
Between Groups	3	703.79	234.60	1.14	n.s.
Within Groups	64	13089.98	204.53		
Total	67[a]	13793.77			

[a]I.Q. scores for two subjects in VTR-CHKL were not available thus accounting for the total of 68 measurements.

CHAPTER VII

INSTRUMENTS AND THEIR CHARACTERISTICS

Recording Equipment

The speeches were recorded and played back on half-inch tapes on a Sony Videorecorder CV2600. Playback was viewed on a nine-inch Sony monitor. Aside from a slight tearing of the image at the top of the screen and the relatively small size of the screen, the pictures produced on this equipment were satisfactory. The television camera was a simple Grundig, with no built-in monitor and no image controls, mounted on a sturdy tripod. The Sony monitor was used during recording to guide the camera work. An Angenieux f2.8 zoom lens was attached. The microphone was a unidirectional Electrovoice Sound Spot model 644 mounted on a microphone stand. A pre-amplifier was used on the microphone line between the microphone and the recorder. The need for a powerful unidirectional microphone was shown in the pilot studies where it was evident that sound quality was going to be much more difficult to control than image quality. The subjects' oral evaluations and oral reactions to self-confrontation were recorded on a Sanyo audio cassette tape recorder. The quality of sound produced in these recordings was poor with many of the gymnasium noises being recorded. Since these recordings were to be used only as an observation aid by the experimenter and not as feedback to the speakers or for evaluation by the judges, their poor quality was not considered detrimental to the experiment.

The pilot studies indicated that the television camera was not a completely objective eye in the hand of the cameraman but that it could be used to accentuate certain elements in the speaker's performance. Since the deliberate seeking out of faults in a speaker's delivery with the camera could influence the speaker's subsequent self-evaluation, a shot sequence to be used with each speaker was developed and followed during the videotape recordings. Basically this involved 1) an opening medium shot of head and upper torso; 2) slow zoom to close-up of face; 3) zoom back and tilt down body to feet; 4) tilt up to long shot of full body; 5) slow zoom-in to opening medium shot and hold. If the speech were a long one the holding of the final shot could be boring. In that case the sequence would be repeated.

Evaluation Instruments

Subjects in groups VTR-CHKL and CHKL used an evaluation form developed for this study called the Speech Ruler. The judges used a revised version of the Speech Ruler. Copies of each may be found in Appendix D and E respectively.

The Speech Ruler consisted of twenty questions, each followed by a scale upon which the subject would place a check-mark in evaluating his speech. The scales were constructed such that the left poles represented a low evaluation and the right poles a high evaluation of performance. All the scales, except two, were divided into five sections. The two exceptions were nine-section scales dealing with the pace and loudness of delivery, elements of performance which could be evaluated lowly in two directions, e.g. "too loud" and "not loud enough,"

with a high evaluation going to a performance falling between the extremes.

No point values were assigned to the scales on either the Speech Ruler or the judges' checklist during the evaluation sessions. It was hoped that the evaluators would view the scales as a continuum. For the purposes of processing and reporting data, point values were attached to the scales.

The judges' checklist was identical in content and scope to the Speech Ruler, with the same number and type of questions. Some of the vocabulary was changed and the detailed instructions deleted. Only the last question (IV b) on the judges' checklist was different from the Speech Ruler. Question IV b on the Speech Ruler asked the speaker to evaluate himself in terms of his previous speaking performances, an evaluation requiring knowledge of the speaker. Since the judges did not know the speakers, their equivalent question was rephrased, and the instructions given the judges brought this to their attention. The judges' question asked how much improvement they would recommend for the speaker based on their experience with pupils of that age group.

Observation Instruments

Three sets of guide sheets were drawn up to guide the experimenter in his observations.

Prior to the delivery of the first speech, each subject gave some background information which was recorded on Data and Observations No. 1 (see Appendix B). Observations made during the recording and playback were kept on Data and Observations No. 2 (see Appendix B). Subjects in groups VTR-CHKL and VTR having five playback sessions each

had five Data and Observations No.2 records kept. Subjects in groups CHKL and CNTL had one such record completed when they were shown the playback of their performances after their last speech.

The verbal reactions and oral evaluations made by each subject were recorded on audio cassettes. These tapes were later subjected to simple content analysis using a form called Oral Evaluation Analysis which may be found in Appendix F. This form reduced verbalized reactions to self-confrontation to a reportable form, a count of the number of evaluative statements made following a speech delivery categorized in terms of their reference to the content of the speech, visual, and verbal elements of delivery, and general, non-specific statements.

CHAPTER VIII

DATA ANALYSIS PLAN

Subjects in treatment groups VTR-CHKL and CHKL used the Speech Ruler in evaluating their own speeches. Each question on the Speech Ruler had been answered by a check-mark placed on the five-sectioned scale. A value of one was given to check-marks at the low end of each scale and a value of five to those at the high end. Values of two, three and four were given for marks falling on sections between the extremes. These values were totalled to reach a subject's total self-evaluation score to be used in the processing of the data. The maximum score possible was 100, that is a score of five on each of the 20 questions. The minimum was 20, that is a value of one on each of the 20 questions. Each subject delivered five speeches so that for each subject in CHKL and VTR-CHKL, five scores were generated. The data reported and processed in this study as "self-evaluations" are these scores.

Subjects in all four treatment groups had each of their video-taped speeches evaluated by three judges. The judges' evaluations based on the speech checklist were transformed into scores in the same manner as with the Speech Ruler. This resulted in three total scores for each of the five speeches delivered by the subject. These three totals were summed and averaged giving a mean judges' score for each speech. The data reported and processed as "judges' evaluations" in this study are, then, the judges' mean scores for each speech by each subject.

The judges' mean scores were subtracted from the subjects'

scores for each speech in groups VTR-CHKL and CHKL. This yielded a "discrepancy score" showing the differences between self-evaluations and judges' evaluations.

The data were tabulated showing for each subject in groups VTR-CHKL and CHKL self-evaluation scores for each speech, judges' evaluations for each speech, and discrepancy scores for each speech. The data for subjects in groups VTR and CNTL were tabulated showing the judges' evaluation for each speech. These data were subjected to simple statistical analysis which produced means and standard deviations for each group. The means were then subjected to single-factor and multiple-classification analyses of variance using formulas found in Hill and Kerber (1967). Since an assumption underlying the analysis of variance is equal variances within classes, tests for homogeneity of variances were conducted. A simple test (Popham, 1967) of homogeneity was applied. This involved calculating the variances of subgroups and dividing the largest by the smallest yielding an F ratio which could be tested for significance using an F table. If the F ratios were found to be significant, the variances were submitted to Bartlett's more powerful test for homogeneity of variance described by Winer (1962). If the variances still proved to be significantly non-homogeneous the analysis of variance was carried out with the fact of non-homogeneity being reported. Glass and Stanley (1970) report that the analysis of variance is a robust test which can withstand some failure to meet its assumptions. They also provide tables indicating the change in values of significance resulting from a lack of homogeneity of variances. If an analysis of variance indicated that a significant difference between means existed the means were submitted

to Scheffé's multiple comparison of means test (Hill and Kerber, 1967) to determine which pairs of means were significantly different.

Single-factor analyses of variance were conducted on the judges' evaluations within each treatment group, and on the self-evaluations and on the discrepancey scores of groups VTR-CHKL and CHKL to determine whether any trend within any of these groups was indicated over the five speech sessions. Multiple-classification analyses of variance were conducted to measure differential effects of experimental treatment on evaluation over the five speeches. The data produced in these analyses are presented in analyses of variance tables.

The oral self-evaluation made by each subject in each group and recorded on audiotape were subjected to simple content analysis. The number of evaluative statements was counted by the experimenter and classified as to whether they referred to speech content, a visual component of delivery, a verbal component or to general non-specific reactions to the speech. The number of evaluative statements by each subject over the five evaluation sessions was tallied. This data was tabulated and subjected to simple statistical analyses which generated means and standard deviations for each treatment group by category of evaluative statement and by total number of evaluative statements. These means were subjected to single-factor analyses of variance in the same manner as described above.

The level of significance chosen for all these analyses was an F ratio with a $p<.05$.

The rest of the data collected and to be presented in this study will be reported in a descriptive form with measurements not reaching above the level of frequencies. Chi-square analyses were conducted on some of these frequencies.

RESULTS AND CONCLUSIONS

CHAPTER IX

RESULTS

The judges' evaluations, the self-evaluations and the discrepancy scores for each subject and for each speech trial are listed in Table 14, Appendix H. The judges' evaluations were scores derived from averaging the three judges' total scores for each speech. the self-evaluation scores were the total scores for each speech by each subject in groups VTR-CHKL and CHKL, that is for those subjects who used the Speech Ruler to evaluate their own speeches. The discrepancy scores refer to the difference between the self-evaluation scores and judges' scores for groups VTR-CHKL and CHKL.

The first two hypotheses tested referred to the improvements in the speech performance of each group over the five speech trials as measured by the judges' evaluations. The judges' evaluations as expressed in Table 14 were submitted to statistical analyses. Mean scores for each speech trial for each treatment group were calculated and are entered in Table 4 along with the standard deviations. Tests for homogeneity of variance were carried out and indicated that the differences in variances within treatment groups and among the four groups were not significant. The same information is displayed in Figure 1, with the standard deviations of VTR-CHKL and CNTL shown as error bars. The error bars for VTR and CHKL were omitted for the sake of neatness.

TABLE 4.--Means of judges' evaluations; standard deviations in parentheses

Speech Trial	Treatment Group							
	CHKL		VTR-CHKL		VTR		CNTL	
	M.	S.D.	M.	S.D.	M.	S.D.	M.	S.D.
1	53.00	(14.26)	59.14	(7.45)	49.29	(13.58)	50.80	(6.54)
2	54.50	(7.52)	59.81	(11.54)	53.14	(10.49)	52.80	(9.05)
3	53.90	(7.14)	62.62	(8.28)	53.36	(9.44)	52.60	(8.95)
4	54.90	(7.62)	64.14	(6.54)	53.43	(10.03)	51.47	(7.74)
5	53.15	(6.87)	63.76	(9.71)	53.64	(8.68)	53.00	(7.73)

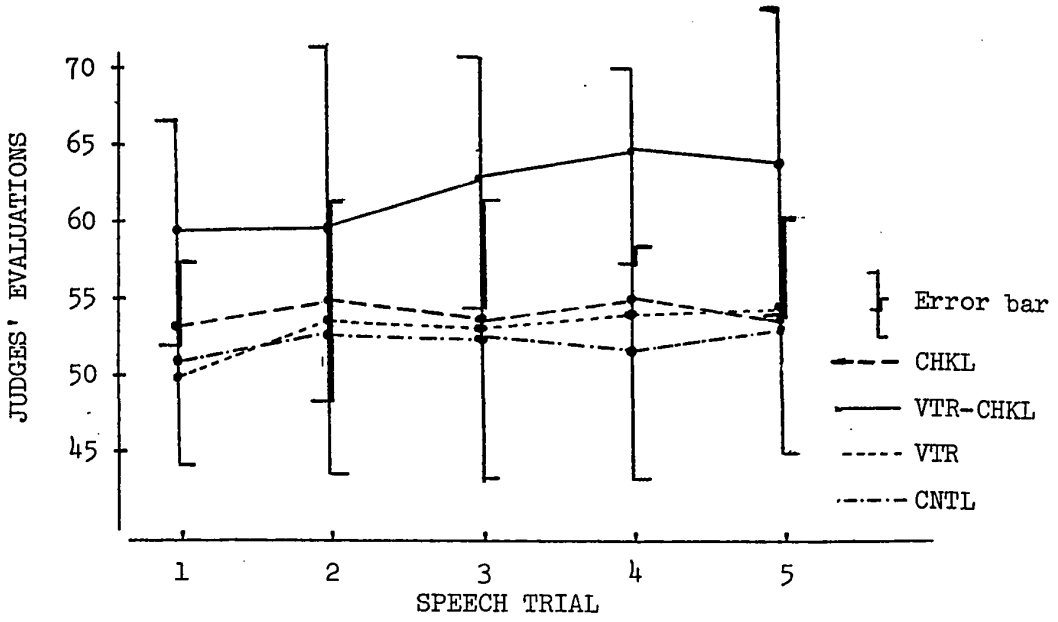


Fig. 1.--Judges' evaluations over five trials; error bars showing standard deviations for VTR-CHKL and CNTL.

H-1 stated that subjects in all four treatment groups would exhibit significant improvements in their speech delivery performances over the five trials as measured by the judges' evaluations. This hypothesis was tested through four statistical hypotheses, one for each group: an analysis of variance of the judges' evaluations will indicate a significant difference in mean scores among the five trials and that difference will be exhibited as an increase in mean scores over the five trials such that

H-1 (CHKL) : $M_5 > M_4 > M_3 > M_2 > M_1$

H-1 (VTR-CHKL) : $M_5 > M_4 > M_3 > M_2 > M_1$

H-1 (VTR) : $M_5 > M_4 > M_3 > M_2 > M_1$

H-1 (CNTL) : $M_5 > M_4 > M_3 > M_2 > M_1$

The results of these analyses of variance may be found in Table 5. The analyses of variance shown in Table 5 indicate that in none of the treatment groups was the change in mean scores over five trials significant enough to reject the null hypothesis at the .05 level. H-1, therefore, could not be accepted and no further tests were undertaken to determine which individual means were greater within any treatment group.

H-2 stated that subjects receiving augmented feedback would exhibit significantly greater increases over five trials than those not receiving such feedback. It was hypothesized that the increases in performance would be in the following descending order: VTR-CHKL showing the largest improvement, then VTR, then CHKL and, finally, CNTL showing the smallest increase. The specific statistical hypothesis tested was: a multiple classification analysis of variance of the five speech

TABLE 5.--Analyses of variance of judges' evaluations within each treatment group

Source of variation	Degrees of freedom	Sum of squares	Mean squares	F	
CHKL (n=20)					
Between means	4	52.90	13.22	< 1	n.s.
Within means	95	5422.10	57.07		
Total	99	5475.00		H_o not rejected	
VTR-CHKL (n=21)					
Between means	4	440.68	110.17	1.33	n.s.
Within means	100	8288.17	82.88		
Total	104	8728.85		H_o not rejected	
VTR (n=14)					
Between means	4	179.90	44.98	< 1	n.s.
Within means	65	7834.45	120.53		
Total	69	8014.35		H_o not rejected	
CNTL (n=15)					
Between means	4	52.99	13.25	< 1	n.s.
Within means	68[a]	5668.14	83.36		
Total	72[a]	5721.13		H_o not rejected	

[a]Speech No. 5 of two Ss not evaluated by judges

trials for the four treatment groups will indicate a significant trial X group interaction effect indicating that the increases in performance of the four treatment groups vary significantly from each other, such that:

H-2 : VTR-CHKL >VTR >CHKL >CNTL

Since it could not be accepted in H-1 that such increases did occur to a statistically significant extent, it became unlikely that the testing of H-2 could indicate differential increases in performance.

The analysis was nevertheless carried out and its findings are summarized in Table 6. The lack of significance in the trial-group variation indicated that H-2 could not be accepted.

TABLE 6.--Multiple-classification analysis of variance of judges' evaluations of the four treatment groups' performance over five trials

Source of variation	Degrees of freedom	Sum of squares	Mean squares	F	
Between groups	3	6068.86	2022.93	24.38	p<.01
Between trials	4	408.00	102.00	1.24	n.s.
Trials X groups	12	318.47	26.20	< 1	n.s.
Within	328	27212.86	82.96		
Total	347	34008.19		H_o not rejected	

The data of Table 6 did indicate, however, that a significant difference existed between the performances of the four groups. Scheffé's test for the multiple comparison of means was applied to the grand means for each group and it was found that this variation resulted from the overall higher performance of group VTR-CHKL over the other three groups. No significant differences were found between the other groups. A look at the means for the four groups as listed in Table 4 and illustrated in Figure 1 confirm this finding that, despite the random assignment of subjects to different treatment groups, group VTR-CHKL was initially superior to the other groups and maintained that superiority throughout the five trials.

The subjects' self-evaluations were submitted to statistical analyses. Subjects in groups CHKL and VTR-CHKL had used the Speech

Ruler to assess their own performances. The data from these two groups may be found in Table 14, Appendix H. The mean scores for each speech trial along with their standard deviations were calculated and are entered in Table 7. The standard deviations were analysed for homogeneity of variance and found not to vary significantly within either group. The distribution of the means over the five trials is shown graphically in Figure 2.

TABLE 7.--Means of self-evaluations; standard deviations in parentheses

Speech Trial	Treatment Group			
	CHKL		VTR-CHKL	
	M	S.D.	M	S.D.
1	70.25	(16.07)	61.90	(11.69)
2	76.50	(13.69)	71.81	(11.62)
3	80.65	(13.79)	78.52	(11.53)
4	80.95	(13.23)	80.33	(11.35)
5	86.15	(10.08)	83.00	(14.61)

H-3 stated that subjects using the self-evaluative checklist would exhibit statistically significant increases in their speech performances over five trials as measured by their own evaluations. The statistical hypothesis tested for each group was: an analysis of variance of the means of each speech trial will indicate a significant difference between the scores on the five trials and that difference will be exhibited as an increase in mean scores over the five trials such that:

H-3 (CHKL): $M_5 > M_4 > M_3 > M_2 > M_1$

H-3 (VTR-CHKL): $M_5 > M_4 > M_3 > M_2 > M_1$

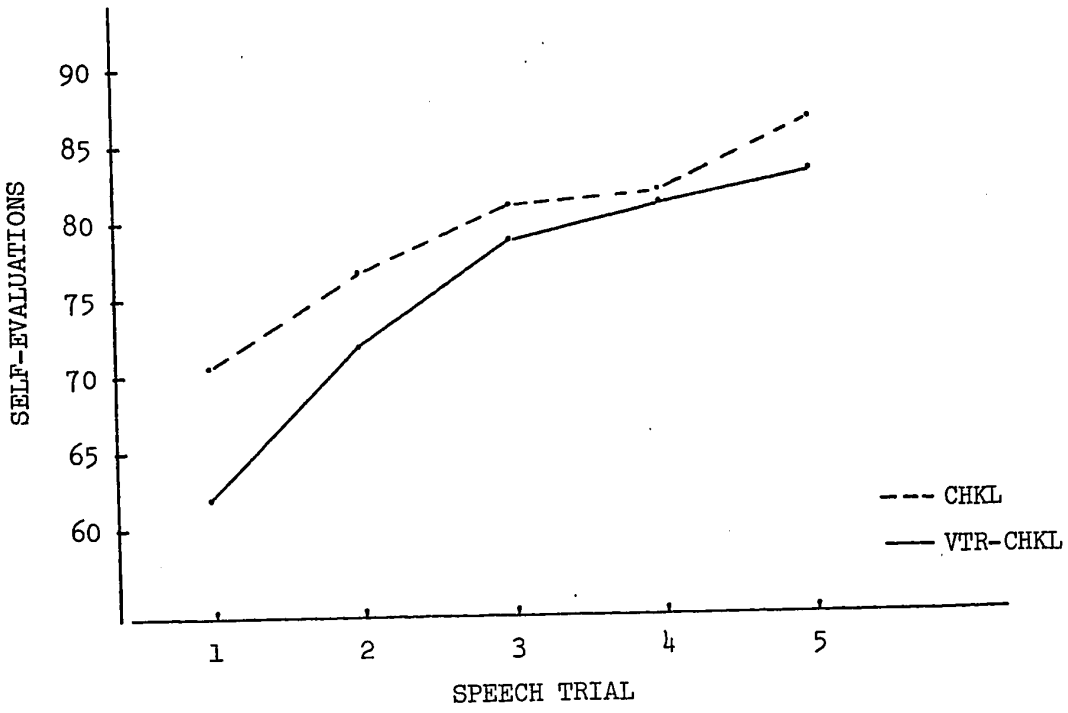


Fig. 2.--Means of self-evaluations over five trials.

The results of these analyses, found in Table 8, indicated that the null hypothesis for both groups could be rejected, the F ratio falling within the level of significance set for this experiment. Scheffé's test for the difference between the means was applied and it was shown that for group CHKL, $M_5 > M_1$ was significant, and that for group VTR-CHKL, $M_3 > M_1$, $M_4 > M_1$, and $M_5 > M_1$ were significant. The differences between the other means, though sometimes high and though all in the direction hypothesized, were not statistically significant. An examination of Table 7 and Figure 2 confirm these findings. H-3, then, was accepted to the extent indicated.

TABLE 8.--Analyses of variance of self-evaluations in groups CHKL and VTR-CHKL

Source of variation	Degrees of Freedom	Sums of squares	Mean squares	F	
CHKL (n=20)					
Between means	4	2001.70	500.43	2.49	$p < .05$
Within means	95	19069.30	200.73		
Total	99	21071.00		H_o rejected	
VTR-CHKL (n=21)					
Between means	4	6015.67	1503.92	9.58	$p < .01$
Within means	100	15684.96	156.85		
Total	104			H_o rejected	

H-4 stated that subjects receiving videotaped playback of their speech delivery would evaluate their performances differently from those not receiving such playback. No direction to this difference between the two groups nor to the interaction between group and trial was hypothesized. Statistically, the hypothesis was: a multiple-classification analysis of variance of the mean self-evaluation scores between the two groups will indicate a significant difference between them, such that:

H-4: VTR-CHKL $\neq$ CHKL

The results of this analysis may be found in Table 9. The analysis indicated that the null hypothesis could be rejected in favour of H-4, that is, subjects in group VTR-CHKL evaluated themselves significantly differently from those in CHKL. An examination of the means indicates that those subjects receiving videotaped playback produced lower self-evaluation. Table 9 further indicates that no interaction effect between trials and groups was evident, that is, no significant

differential gains between the two groups in the self-evaluations occurred over the five trials.

TABLE 9.--Multiple-classification analysis of variance of groups CHKL and VTR-CHKL's self-evaluations over five trials

Source of variation	Degrees of freedom	Sums of squares	Mean square	F	
Between groups	1	734.06	734.06	12.15	$p < .01$
Between trials	4	8467.28	2116.82	4.21	$p < .01$
Trials X groups	4	356.59	89.14	< 1	n.s.
Within	195	33948.76	174.09		
Total	204	43506.69		H_o rejected	

H-5 stated that subjects receiving augmented feedback in the form of videotaped playback of their speech deliveries would produce self-evaluations closer to the judges' evaluations than those not receiving such feedback. The criterion measure was the scores produced by subtracting the judges' evaluations from the self-evaluations. These discrepancy scores may be found in Table 14, Appendix H. The means and standard deviations for these discrepancy scores were calculated and entered in Table 10 and Figure 3.

The statistical hypothesis tested was: a multiple-classification analysis of variance of the mean discrepancy scores in groups CHKL and VTR-CHKL will indicate a significant difference between the scores of the two groups such that:

H-5: CHKL > VTR-CHKL

TABLE 10.--Means of discrepancy scores; standard deviations in parentheses

Speech Trial	Treatment Group			
	CHKL		VTR-CHKL	
	M	S.D.	M	S.D.
1	17.25	(19.73)	3.04	(10.23)
2	22.00	(16.35)	11.90	(12.14)
3	26.75	(13.57)	15.90	(13.39)
4	26.15	(14.90)	16.19	(12.26)
5	33.00	(10.26)	19.23	(12.71)

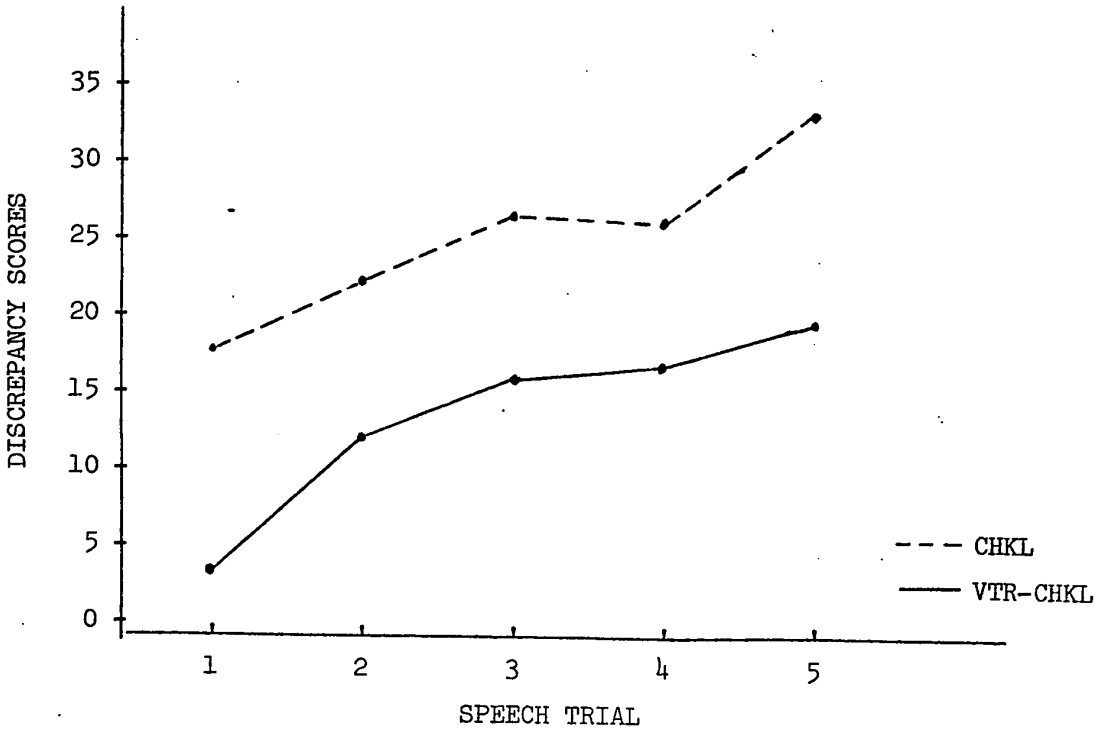


Fig. 3.--Means of discrepancy scores over five trials.

The results of this analysis may be found in Table 11. The analysis indicated that the null hypothesis could be rejected, thus confirming H-5 that the discrepancy scores for CHKL were larger than those for VTR-CHKL. Table 11 further indicates that the change in discrepancy scores for both groups over the five trials was significant but that no interaction effect between the two groups was evident. An examination of Figure 3 confirms these findings.

TABLE 11.--Multiple-classification analysis of variance of the discrepancy scores of groups CHKL and VTR-CHKL

Source of variation	Degrees of freedom	Sum of squares	Mean square	F	
Between groups	1	7137.03	7137.03	37.54	$p < .01$
Between trials	4	5920.95	1480.23	7.78	$p < .01$
Trials X groups	4	134.34	33.58	< 1	n.s.
Within	195	37068.68	190.09		
Total	204	50261.00		H_o rejected	

H-6 stated that subjects receiving videotaped feedback would perceive more in their speech performances, as measured by an analysis of their oral self-evaluations, than those not receiving such feedback. Subjects in all four treatment groups evaluated their speeches orally, immediately after delivery in CHKL and CNTL and immediately after video playback in VTR and VTR-CHKL. These oral evaluations were submitted to a simple content analysis using the Oral Evaluation Analysis form of Appendix F. The mean number of self-evaluative

statements for the five speech trials in each group are listed in Table 12 according to the type of statement. The same data is represented in a bar graph in Figure 4.

TABLE 12.--Mean number of oral self-evaluative statements by subjects in four treatment groups; standard deviations in parentheses

Group	Type of Statement				
	Content	Visual	Verbal	General	Total
CHKL (n = 19)	3.63 (2.11)	3.15 (2.87)	3.89 (2.92)	4.42 (1.99)	15.11 (5.69)
VTR-CHKL (n = 20)	4.25 (1.86)	4.85 (2.89)	3.95 (1.73)	4.65 (2.12)	17.70 (4.49)
VTR (n = 14)	3.35 (1.40)	4.07 (2.95)	4.50 (2.98)	4.64 (1.84)	16.57 (6.03)
CNTL (n = 15)	3.06 (2.34)	1.00 (2.33)	1.80 (1.70)	4.73 (1.94)	10.60 (4.45)

The statistical hypothesis tested was: an analysis of variance of the number of self-evaluative statements will indicate that a significant difference exists among the four groups in each of the four categories and in the total number of statements, such that:

H-6: VTR-CHKL > VTR > CHKL > CNTL

The results of these analyses were entered in Table 13. It was found that H-6 could be partially accepted. Scheffé's test was applied and it was found that VTR-CHKL > CNTL was significant in the visual and total categories and VTR > CNTL was significant in the visual, verbal and total categories.

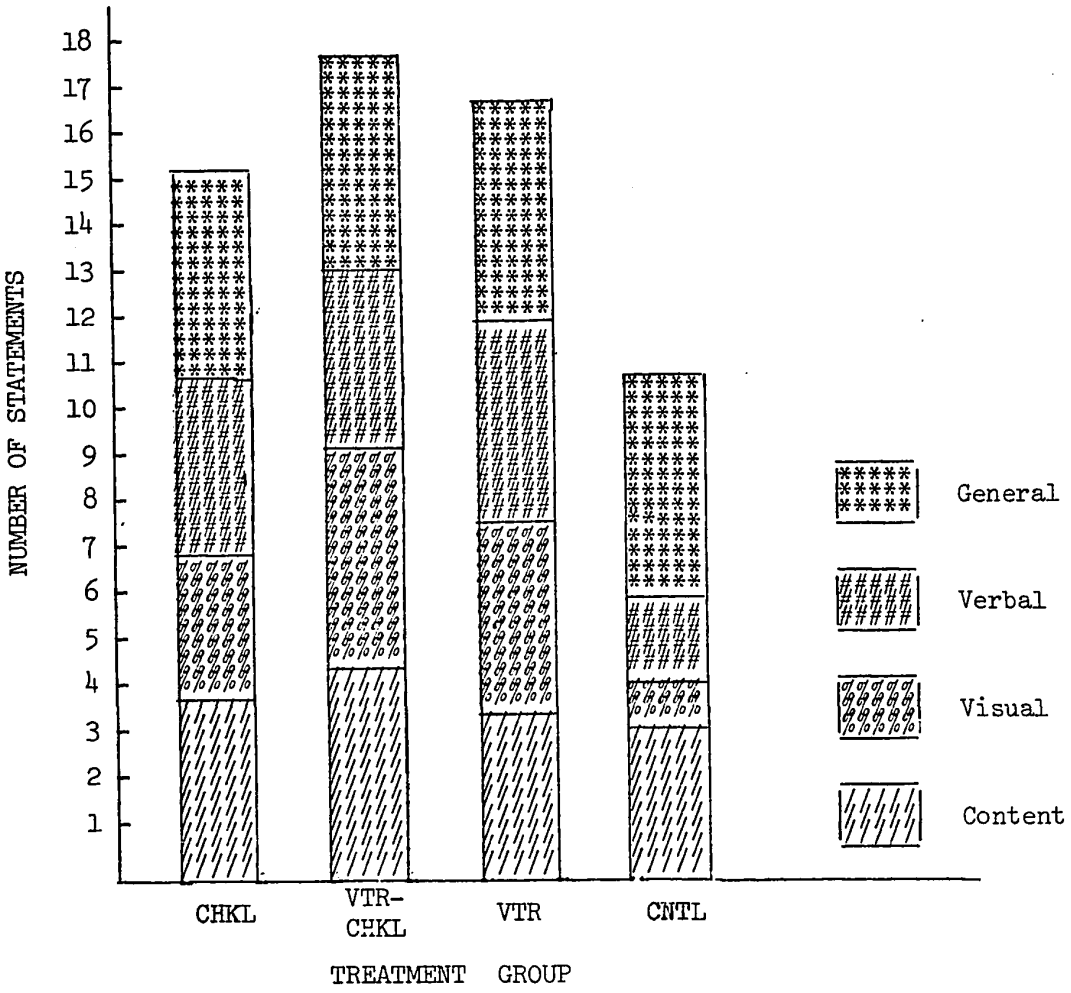


Fig. 4.--Mean number of oral self-evaluative statements in four treatment groups.

TABLE 13.--Analyses of variance of the number of oral self-evaluations among four treatment groups

Category of Statement	Source of variation	Degrees of freedom	Sums of squares	Mean square	F	
Content	Between groups	3	13.47	4.49	1.32	n.s.
	Within	64	248.34	3.88		
	Total	67	261.81			
Visual	Between groups	3	135.80	45.26	5.84	$p < .01$
	Within	64	496.01	7.75		
	Total	67	631.81		H_o rejected	
Verbal	Between groups	3	64.00	21.33	3.72	$p < .05$
	Within	64	366.64	5.72		
	Total	67	430.64		H_o rejected	
General	Between groups	3	0.93	0.31	< 1	
	Within	64	117.35	1.83		
	Total	67	118.28			
Total	Between groups	3	468.87	156.29	5.83	$p < .01$
	Within	64	1715.02	26.79		
	Total	67	2183.89		H_o rejected	

In addition to the testing of the above hypotheses in a statistical mode further data was collected in response to two questions asked in this study.

Q-1 asked: Do elementary school children identify with the video and audio images of themselves presented during videotaped playback? All seventy subjects were given the opportunity of seeing the videotaped recordings of their speech delivery performance. Subjects in VTR-CHKL and VTR received this information as part of the experimental treatment; subjects in CHKL and CNTL received it once the experimental session was

over.

Immediately after viewing themselves, the subjects were asked: a) "How did you look on TV? Did you look as you usually look or did you look different?" and b) "How did you sound? Did you sound the same as you usually do or did you sound different?" The following frequencies of response were given:

Look like self	Yes	55
	No	15
Sound like self	Yes	15
	No	55

Twenty-three subjects offered no further comment beyond the affirmative or negative for the question dealing with looks. Of those who did comment further, most placed negative conditions upon their affirmative answers, these comments taking the general form of "Yes, but . . . ." Nine thought they looked older while five thought they looked younger on television. Fourteen thought they looked "horrible", "ugly", "like Frankenstein", "funny", or used some similar negative term. Six commented on their hair, two on some other facial characteristic. Four of those who responded negatively thought they looked like somebody else.

The reaction to voice fidelity was the reverse, in frequency, to that of video fidelity. The accompanying comments, however, were not as intense. Several subjects mentioned that they had experience with audiotapes. Fifteen made no comment beyond the simple affirmative or negative. Twelve found that they sounded "younger", "higher", or "squeaky". Four claimed that they sounded "older". Eight said they

sounded "hoarse". Five thought they sounded like somebody else.

Q-2 asked: How do elementary school children react during the playback of a videotaped presentation of themselves? The subjects were observed closely by the experimenter during playback. The experimenter's observations were guided by the categories set down in the Data and Observations No. 2 form of Appendix B. Five categories of observation were used: 1) subject's distance from the screen, 2) facial expression, 3) hand movement, 4) gross bodily movement or position, and 5) object of eye focus.

It was found in the pilot studies that students inched up closer and closer to the television monitor during playback. This movement was not replicated in the experimental sessions. The most obvious explanation for this is that during these sessions, the subjects were seated and a desk was placed between them and the monitor, whereas in the classroom, students were generally seated on the classroom floor with no obstruction, other than other students, between them and the rather small nine-inch monitor. In the classroom there was pressure from those seated behind on the ones up front. In the experimental sessions there was no crowding factor.

Facial expression during playback ran the whole gamut from deadpan to the melodramatic overacting of expressions of horror. One of the most common expressions, and the one that seemed the most genuine, was one which is difficult to express in words. It was one which occurred frequently in the first moments of self-confrontation and which can be described neither as joy nor pain, nor fear, but appearing to carry traces of each of these emotions. The word which best describes it is "open". The eyes were kept wide open, focused intently on the

monitor screen; the mouth slightly open or about to open; there was a slight leaning forward of the face and body. It was an expression approaching awe yet without the sense of being overwhelmed. The subject seemed oblivious to his surroundings and seemed to be judging himself intently. It seemed like an expression one might make while judging and examining oneself in a mirror when nobody else is looking. For some subjects this expression lasted only a moment and was quickly converted into some overt, communicable expression such as one of disgust with dramatic facial contortions or of embarassment with a turning away from the screen. A few subjects maintained this "open" expression throughout the video replay.

Other common facial movements were smiles, lip biting, and expressions of concentration. Less prevalent but interesting were certain lip movements: some subjects were observed repeating or rehearsing their speeches in unison with their videotapes, particularly after the second or third delivery.

The hands were generally quite active during playback except during those moments of intense concentration described above as "open". Most of the hand movements were face-related, that is, they tended to flutter around the face pinching, pulling and scratching, or were used as supports to rest the chin and cheeks on. The hands were also often used to play with some object, usually the pencil on the desk.

There seemed to be no fixed or significant bodily postures or movements. During the initial moments of videotaped playback the subjects' bodies tended to be tense and leaned forward towards the screen. Some subjects maintained this alert position during playback but most assumed more comfortable positions.

Eye focus tended to be on the screen but there was much variability here. A small number of subjects took an initial glance at the screen at the beginning of the playback but they consistently avoided it and looked at the floor, the desk, their hands, anywhere but at the screen. The majority kept their eyes on the screen except for quick glimpses at the experimenter to see his expression.

CHAPTER X

DISCUSSION OF RESULTS

It was hypothesized that improvements in speech delivery performance measured by increases in judges' evaluations over five speech trials would occur for all four treatment groups (H-1) and that these increases would be greater for those subjects receiving augmented feedback about their performances (H-2). Subjects in CHKL and VTR-CHKL scored highest on their fourth delivery while those in VTR and CNTL scored highest on their fifth. Subtracting the highest mean score for each group from its score on the first trial, we find the following increases: CHKL = 1.90, VTR-CHKL = 5.00, VTR = 4.35, CNTL = 2.20. The hypothesized improvements, then, did occur and in the direction hypothesized (except for the reversed positions of CHKL and CNTL) but not in statistically significant ($p < .05$) quantities.

Accepting the hypotheses to be still theoretically sound we must look at the particular experiment for signs of weakness which may explain why they could not be accepted in this study. A possible factor militating against an improvement in performance over five trials was one of overloading built into the experimental design. Speech delivery involves complex psychomotor skills. In a period of less than two hours, the subjects delivered and evaluated five speeches as well as took part in the novel experience of videotaped self-confrontation. It is possible that this system was overloaded and that

measurable changes in performance could not be exhibited in the short-run-time involved here. The same experimental variables, manipulated in the same manner, with the same subjects, but spaced out over a longer period of time might have resulted in greated changes in performance. As we saw in the Review of the Literature, Ward (1970) found that placing a one-week delay between re-teach sessions in a micro teaching experiment produced measurable changes in performance not found when the re-teach session followed immediately after feedback. Since Ward's study appeared only after the data for this study had been collected, his experience could not be incorporated here.

Another factor complicating the results of this experiment was the initial superiority of the subjects in group VTR-CHKL. Despite the random assignment of subjects to the various treatment groups, despite the fact that the judges were not aware of which groups the subjects had been in, and despite the fact that all subjects had received identical instructions prior to the delivery of their first speech, the subjects in VTR-CHKL were judged to be superior to the other groups on their first speech and maintained this superiority throughout the five trials. The greater improvement over the five trials which this group exhibited, even though not statistically significant, may not, then, be attributed solely to the different experimental treatment they received but may also be due to some unknown factor which caused the initial superiority. A multiple-classification analysis of covariance, using the scores of the first trial as a covariant against which to correct the means of the other trials, would be a more powerful tool than the analysis of variance used here and is recommended for future research. Since initial differences between the

groups were not anticipated and since the analyses of variance did not indicate any significant increases in performance within any treatment group which could be overruled by an analysis of covariance, that analysis was not used here.

Looking at the subjects' performances as seen through their own evaluations, it was hypothesized that the subjects would see their performances improving over the five trials (H-3), that those receiving videotaped feedback would evaluate their performances differently from those not receiving such feedback (H-4), and that the differences between self- and judges' evaluations would be smaller for the video feedback groups (H-5). The results of this experiment confirmed these three hypotheses. Subjects' mean self-evaluations increased steadily over the five trials with an increase of 15.90 for CHKL and 21.10 for VTR-CHKL from trial one to trial five. The self-evaluations of subjects receiving video feedback were lower on each of the five trials than of those not receiving such feedback. The combination of this lower self-evaluated score with the higher judges' evaluations for this group resulted in a lower discrepancy between self- and judges' evaluation for VRT-CHKL.

From a cybernetic point of view, then, the data seem to confirm our hypothesis that subjects receiving video feedback, having more information upon which to base their evaluations, are less dependent upon subjective recall of a past performance and are more capable of seeing themselves as others see them. If we accept the judges' evaluations as objective observations, at least in the sense that the same judges viewed all the performances, used the same evaluative tool, and had their scores averaged together to form one score,then it can be argued that the lower self-evaluations

generated in the VTR-CHKL group, being closer to the judges' evaluations, were more objective.

Two factors weaken this argument, however. The first is the indication found in the analysis of the judges' evaluations that the subjects in VTR-CHKL might differ in some important factor from those in CHKL. Whatever factor was operating to make this group superior in the judges' evaluation of their speech delivery performance may be the same factor operating to lower the scores in the self-evaluations. There may be, for example, a character type which predominated in this group despite attempts to randomize the assignment of subjects to different groups. The second factor is the evidence of the large increases in self-evaluated performance over the five trials shown in group VTR-CHKL. The subjects in VTR-CHKL although they evaluated themselves lower than did those in CHKL, did view their performances as improving significantly (Table 8) over the five trials at a rate greater (not significantly, Table 9) than those in CHKL. A look at the slopes of the curves in Figure 2 confirms this. If video feedback were operating to produce a more objective self-evaluation, one closer to the judges' evaluation, then one might expect a self-evaluation curve similar to that of the judges who recorded a slight but not significant improvement over the five trials. It should be pointed out here, however, that the judges did not experience the performances in the same sequence as did the subjects. The subjects, of necessity, saw and evaluated their performances in a sequential order from the first trial through to the fifth, whereas the judges saw a randomized presentation of performances. The subjects could operate under a bias that their speeches would improve with repetition whereas the judges,

even if they held this bias, could not put it into effect not knowing which of a subject's five speeches they were viewing. All that can be safely concluded from these results is that the judges evaluated the performances lower than did the subjects and that, while the subjects receiving video feedback in this experiment could be said to evaluate themselves more objectively than did those not receiving such feedback, their perceived rate of improvement over five trials was greater indicating the possible presence of some factor counteracting the pressure towards objectivity produced by video feedback.

The Speech Ruler provided the researcher with numerical data to measure changes in self-evaluation and was useful to the subjects as an evaluative guideline. Nevertheless, it did involve to a certain extent an external and artificial element in the self-evaluative process in that subjects using it had to respond to each item whether meaningful or not to them. The oral evaluations which subjects made of their performances, on the other hand, were not structured or regulated by the experimenter. They proceeded from the subject and involved the construction, or verbalization, of an evaluation rather than the response to a form. The greater number of these oral self-evaluative statements (Table 12) generated by those subjects in VTR-CHKL and VTR confirmed in part the hypothesis (H-6) that subjects receiving video feedback (VTR-CHKL and VTR) perceived more in their performance than those not receiving such feedback (CNTL). The lack of a significant difference between these two groups and group CHKL indicates, however, that offering the performer guidelines with which to focus his attention on his performance may be a sufficient alternative to the use of the videotape recorder.

Finally, this study attempted to describe the subjects' reactions to videotaped self-confrontation. It was found that most subjects recognized themselves visually on the monitor but reacted in a manner characteristic to audiotapes in not recognizing their own voices. The cosmetic effect typical of a mirror was present as evidenced by the intense attention paid by subjects to the images of themselves on the monitor, particularly on initial self-confrontation. The data collected for this part of the study suffered, however, by being based largely on the subjective impressions of the experimenter. More rigid observation schedules would be recommended for future research. The data presented here should be seen as preliminary indications that a study of the learner in interaction with his videotape-based image could constitute a fruitful research area.

CHAPTER XI

CONCLUSION AND RECOMMENDATIONS

The purpose of this study, as set forth in the Introduction, was to evaluate the effectiveness of using television as a learning and evaluation aid rather than as an instructional aid, to use the videotape recorder as a source of feedback which could offer the learner information about his past performances, information which he could evaluate and translate into improved subsequent performances. The processes studied here were self-evaluation and speech delivery. The videotape recorder and the self-evaluative checklist were the intervening variables, the former offering augmented feedback to the subject and a record of the performances to the investigator; the latter offering the subject a framework with which to guide his evaluations and the researcher a measure of self-evaluation.

The mixed results produced by this experiment suggest the need for more research. On the one hand, the absence of significant improvements in performance did not correspond with the expectation of an improvement with repetition or of more improvement for those receiving more feedback. Yet the fact that the results tended in the direction hypothesized suggests that a repetition of the experiment with certain modifications might yield clearer results. The evidence produced by the measures of self-evaluation confirmed in part the hypothesis that augmented feedback would result in more objective self-evaluation. The large gap, however, between self-

and other-evaluation offered evidence of the differences between these two types of evaluation, and, in the light of Nuttin and Greenwald's (1968) findings on the influence of other-evaluation on the learner's feeling of success or failure, suggests the need for further investigation of self-evaluation as an alternative to other-evaluation. The evidence that subjects receiving augmented feedback produced more oral self-evaluative statements gave further indication that the manipulated variables were having some effect. Finally, the investigator's observations concerning the subjects' reactions to video self-confrontation indicate the potential of this medium in affecting behaviour.

The experience gained in carrying out this piece of research points to several further research possibilities, some involving changes in the experimental design used here, others suggesting further areas to be investigated.

Presenting the subject with more time between repeated performances than was provided in this experiment might result in greater improvements in performance. It might be fruitful to repeat the experiment leaving all variables the same but varying the time interval between speech deliveries. One possible design could allow for one hour, one day and one week intervals. This pattern could indicate whether the interval between repeated performances is an important variable and, if so, whether optimal interval patterns exist. It is likely that if such a pattern exists it is in turn influenced by other factors such as the complexity of the task, the performers' personal characteristics, and the nature of the performers' experiences between trials.

The effect that time plays in the judges' evaluations could also be manipulated and made the subject of further research. The videotape recorder allows the evaluator to manipulate time. The tape may be stopped and re-run allowing the evaluator to review the performance and revise his evaluation. Not only can the performance be stopped and repeated in this manner, but it can also, with the newer recorders, be slowed down and stretched out through a slow-motion mechanism. The effect of time can also be modified, as it was in the present study, by reorganizing or scrambling the order of the performances presented to the judges. A scrambled presentation of the videotaped speeches was made in this study so as to overcome the possible bias that a subject's performance would tend to improve with successive repetitions. Another study might undertake to present the same speeches in their original sequence thus allowing the judges to operate under such a bias, and, by a comparison with the results of the present study, to determine the strength of such a bias if it exists. In this way, the videotape recorder's ability to scramble presentations might be put to use as a research tool to investigate the halo effect in evaluations of succeeding presentations. The scrambled presentations could also be exploited in educational evaluations where the sequence of presentation is a factor affecting objectivity. There is a movement in education away from comparing students with each other towards comparing a student's performance with his past performances. The videotape recorder, with its ability to record past performances and play them back in a scrambled sequence, could provide the medium for such an evaluation technology.

Another possible research design, one closer to the reality of

the classroom, would allow the subject access to the evaluations of his teachers and peers. In the present experiment the subject was isolated from these sources of feedback. This was done so as to study self-evaluation in as uncontaminated a manner as possible. This suggested alternate multi-factor design could measure the influence of other-evaluation on self-evaluation. Personality and sociometric factors would most likely have to be accounted for in this design.

Research into the personal characteristics of subjects might indicate that different types react differently to video feedback. Subjects with differing self-images might produce differing self-evaluations. Some persons might not identify with their video images or might do so differently from others. There was some indication of this in the subjects' oral reactions to their video selves. Some were obviously pleased while others were embarrassed. A measure of narcissism correlated to self-evaluations might produce interesting results. It was seen in the literature on self-evaluation (Russel, 1953) that males differed from females and that subjects of high intelligence differed from those of low intelligence in evaluating their own performances. It is also likely that subjects with past experience or training in a task would evaluate themselves differently from those without such experiences. Since data existed on sex, intelligence, and the previous speech training of the subjects in this study, an analysis of this data was undertaken to give some indication of what future research might show. This data is collected in Appendix I. Briefly, it indicated that males, subjects below the median in I.Q. and subjects with no speech training scored lower than females, those above the median, and subjects with previous speech training respectively. No differences

were found between male and female and between low I.Q. and high I.Q. subjects in self-evaluations. Subjects with no previous speech training receiving video feedback evaluated themselves lower than did those with training. This difference was not found for subjects not receiving video feedback. The numbers of subjects in any of these sub-groupings was too small to make any generalizations at this point. The results do indicate, however, that an analysis of the subjects' characteristics would be a fruitful area for further research in the field of video feedback.

A final area of further research suggested by the present study concerns that of training in self-evaluation. The learner, in order to control and alter his future performances, must first be able to perceive what is relevant in the information fed back to him. Then he must be able to analyze it and compare it to a standard of performance. The mere playing back of a videotape does not constitute feedback. The video information must be transformed by the self-regulating organism into some meaningful form. Research could be carried out to determine if and how individuals differ in their abilities to perceive, analyze and evaluate performances. At the same time research could be carried out into the methods of improving evaluating skills. Possible methods include the use of model performances in evaluation execises, the comparison of learner evaluations to expert evaluations, the use of checklists or guides in self- and other-evaluation, and the use of the videotape recorder, not as a source of feedback, but as a reference against which to check and compare self- and other-evaluations.

The introduction of a new technology is often met with fanfare

and promises of revolutionary changes although it is most often immediately used to achieve traditional goals. The videotape recorder was developed to give television broadcasters an electronic means of recording programmes. This remains the primary use of the video recorder. The development of the inexpensive half-inch recorder placed this technology within the reach of smaller institutions and individuals. In education the primary use of the videotape recorder is the recording of programming, of information external to the audience. Schools often record programmes off-air for later presentation to students. Some schools produce their own programmes, some even use students in the production. If these were the only tasks the videotape recorders could accomplish, they would be sufficient to warrant their purchase and utilization. As we have seen in this thesis, however, there is a further educational use to which the videotape recorder can be put. The videotape recorder is unique among the media in its capability to record audio and video information for immediate playback. Exploiting this capability allows the educator to provide the learner with an additional source of feedback, one which invites learner self-evaluation. The utilization of the videotape recorder as an aid for self-evaluation is relatively recent and, while accompanied by much enthusiasm, is still not well tested. The research presented in this thesis, with its mixed results, indicates the necessity for further research in this area.

LIST OF REFERENCES

Ahmann, J., and Glock, Marvin D. 1967. Evaluating pupil growth: principles of tests and measurements. 3rd ed. Boston: Allyn and Bacon.

Alger, Ian, and Hogan, Peter. 1969. Enduring effects of videotape playback experience in family and marital relationships. American Journal of Orthopsychiatry. 39: 86-98.

Alkire, Armand. 1969. The use of videotaped playback in the fields of education and mental health. AV Communication Review. 17,2: 182-200.

Ammons, R.B. 1956. Effects of knowledge of performance: a survey and tentative formulation. The Journal of General Psychology. 54: 279-299.

Annett, John. 1969. Feedback and human behaviour. Harmondsworth, Middlesex, England: Penguin Books.

Becker, Samuel L., Bowers, John W., and Gronbeck, B.E. 1968. Videotape in teaching discussion. Speech Trainer. 17,2: 104-106.

Bloom, Benjamin S., Hastings, J. Thomas, and Madaus, George F. 1971. Handbook on formative and summative evaluation of student learning. New York: McGraw-Hill.

Boyd, Harry S., and Sisney, Vernon V. 1967. Immediate self-image confrontation and changes in self-concept. Journal of Consulting Psychology. 31,3: 291-294.

Cooper, James A., and Allen, Dwight W. 1970. Microteaching: history and present status. Washington, D.C.: ERIC Clearinghouse on Teacher Education.

De Bacy, Diane. 1970. Effect of viewing video tapes of a sport skill performed by self and others on self-assessment. The Research Quarterly. 41,1: 27-31.

Dieker, Richard J., Crane, Loren, and Brown, Charles T. 1968. Repeated self-viewings on closed-circuit television as it affects changes in students' awareness of themselves as speakers. Final report. Kalamazoo, Michigan: Western Michigan University.

Frandsen, Kenneth D., Larson, Carl E., and Knapp, Mark L. 1968. Simulation and self-confrontation in interpersonal communication. Educational Broadcasting Review. 2,2: 18-23.

Glass, Gene V., and Stanley, J.C. 1970. Statistical methods in education and psychology. Englewood Cliffs, New Jersey: Prentice-Hall.

Hagen, Elizabeth P., and Thorndike, Robert L. 1960. Evaluation. Encyclopedia of Educational Research. Chester W. Harris, ed. 3rd ed. New York: Macmillan and Co. 482-485.

Henderson, John. 1964. Using mirror TV to teach speaking. The NAEB Journal. 23: 53-56.

Heinich, Robert. 1968. Systems engineering of education II: application of systems thinking to instruction. Los Angeles: Education and Training Consultants Co.

Hill, Joseph E., and Kerber, August. 1967. Models, methods and analytical procedures in educational research. Detroit, Michigan: Wayne State University Press.

Hirschfeld, Adeline G. 1968. Videotape recordings for self-analysis in the speech classroom. The Speech Teacher. 17,21: 116-118.

Hurley, Carl Edward. 1971. The effects of feedback on psychomotor performance of fourth and sixth grade students. Unpublished Ed.D. dissertation. Columbia, Missouri: University of Missouri.

Jensen, Paul H. 1968. A study of self-evaluation applied to in-service education. Washington, D.C.: U.S. Department of Health, Education and Welfare.

Kagan, Norman, Krathwohl, David R., and Miller, Ralph. 1963. Simulated recall in therapy using videotape- a case study. Journal of Clinical Psychology. 10,3: 237-243.

Merrill, M. David. 1968. Components of a cybernetic instructional system. Educational Technology. 8,7: 5-10.

Mulac, Anthony John. 1968. An experimental study of the relative pedagogical effectiveness of videotape and audiotape playback of student speeches for self-analysis in a basic speech course. Washington, D.C.: U.S. Department of Health, Education and Welfare.

Nelson, Harold E. 1968. Instructional uses of videotape: a symposium. 1. Videotaping the speech course. The Speech Teacher. 17,2: 101-103.

Nielsen, G. 1962. Studies in self-confrontation. Viewing a sound motion picture of self and another person in a stressful dyadic interaction. Copenhagen: Munksgaard.

Nuttin, Joseph, and Greenwald, Anthony G. 1968. Reward and punishment in human learning. Elements of behavior theory. New York: Academic Press.

Peavy, R. Vance. 1969. Cybernetic educational design: an example. Journal of Educational Thought. 3, April: 15-19.

Popham, W. James. 1967. Educational statistics. Use and interpretation. New York: Harper and Row.

Russel, David H. 1953. What does research say about self-evaluation? Journal of Educational Research. 46,8: 561-573.

Robb, Margaret, and Teeple, Janet. 1969. Videotape and skill learning; an exploratory study. Educational Technology. 9,11: 79-82.

Ryan, Paul. 1968. Videotape...thinking about a medium. Media and Methods. 5,4: 36-41.

Ryan, Paul. 1969. Videotape and special education. Audiovisual Instruction. 14,9: 30.

Sawin, Enoch I. 1969. Evaluation and the work of the teacher. Belmont, Calif.: Wadsworth Publishing Co.

Shively, Joe E., Van Monframs, Adrian P., and Reed, Cheryl L. 1970. The effect of mode of feedback in microteaching. Paper presented at the Annual Meeting of the American Educational Research Association. Minneapolis.

Smith, Karl U., and Smith, Margaret Foltz. 1966. Cybernetic principles of learning and educational design. New York: Holt, Rinehart and Winston.

Solomen, Gabriel, and McDonald, Frederick J. 1968. Pre-and post-test reactions to self-viewing one's performance on videotape. Paper presented to the Annual Convention of the American Psychological Association, August 30-September 3.

Stroth, Thomas F. 1969. The uses of videotape in training and development. New York: American Management Association.

Stoller, F.H. 1967. The long weekend. Psychology Today. 1,7: 29-33.

Ward, Phillip M. 1970. The use of the portable videotape recorder in helping teachers self-evaluate their teaching behaviour. California: University of California at Berkeley.

Wert, James E., Neidt, Charles O., and Ahmann, J. Stanley. 1954. Statistical methods in education and psychological research. New York: Appleton-Century Crofts.

Winer, B.J. 1962. Statistical principles in experimental design. New York: McGraw-Hill.

APPENDICES

APPENDIX A

HANDOUT OF GENERAL INFORMATION

Each pupil in the intermediate level population was given a handout summarizing the general instructions given orally by the experimenter prior to the commencement of the experimental sessions. This handout is reproduced below.

> Mr. Jones will be making television recordings of intermediate students giving short speeches. These will start on Tuesday, May 19, and will continue for about one month. Be ready to give a speech of about one minute on one of these subjects:
>
> 1. Pollution
> 2. My Favourite Game
> 3. My Sports Hero
> 4. My Best Friend
> 5. Summer Vacations
>
> If you are chosen to be on television you will be told at least one week in advance. BE PREPARED.

APPENDIX B

DATA AND OBSERVATION FORMS

Background information and observations made by the experimenter during the experiment were kept on two forms. Data and Observations No.1 was used prior to the delivery of the first speech. One copy of this form was collected for each subject. Data and Observations No. 2 was used immediately prior to and during playback of the videotapes to the subjects. There were, then, five copies collected for each subject in groups VTR-CHKL and VTR, corresponding to his or her five playback sessions, and one for each subject in CHKL and CNTL corresponding to his or her one session of playback at the end of the experimental session.

Data and Observations No. 1

1. Name:_____ Code:____ ____ ____ ____ ____

2. Age:_____ _____ 3. Sex:_____ 4. Grade:_____

5. Lorge-Thorndike: V_____ Non-V_____ Composite_____

6. CTBS V_____ R_____ L_____ C_____

7. Videotaped before: No_____ Yes_____

8. Television before: No_____ Yes_____

9. Home movies before: No_____ Yes_____

10. Speech training: No_____ Yes_____

_____

11. Speech experience:_____

_____

12. This speech. Topic:_____

Preparation:_____

13. Tone prior to first session:

lethargic	1	2	3	4	5	hyperactive
apprehensive	1	2	3	4	5	assured
shy	1	2	3	4	5	open
resistant	1	2	3	4	5	accepting
nervous	1	2	3	4	5	calm
bored	1	2	3	4	5	eager

Other:_____

14. Rapport:

poor	1	2	3	4	5	good

15. Notes:

Data and Observations No. 2

Code____ ____ ____ ____ ____

16. Tape_____ Footage_____ to_____

17. Speech time_____ seconds

18. Disturbances: 1. not serious 5. very serious

Speech_____

Playback_____

Evaluation_____

19. Tone prior to playback:

lethargic	1	2	3	4	5	hyperactive
apprehensive	1	2	3	4	5	assured
shy	1	2	3	4	5	open
resistant	1	2	3	4	5	accepting
nervous	1	2	3	4	5	calm
bored	1	2	3	4	5	eager

Other:_____

20. Actions during playback:

a) distance to screen: Same____ Moved____ closer, away

b) face: 1._____

2._____

3._____

c) hands: 1._____

2._____

3._____

d) body: 1._____

2._____

3._____

e) eyes: 1._____

2._____

3._____

21. Oral Evaluation Audio tape no.____ ft.____

a) How did you look on TV? Did you look the same as you usually look or did you look different?

b) How did you sound? Did you sound the same as you usually do or did you sound different?

c) How was your speech? Did you like it? If you had to do it over, is there anything that you would not do again? Is there anything that you would keep the same?

22. Notes (over)

APPENDIX C

ORAL INSTRUCTIONS ON HOW TO USE THE SPEECH RULER

Each subject in groups VTR-CHKL and CHKL was given detailed instructions in the use of the Speech Ruler after delivering his first speech. These instructions took the form of a general introduction by the experimenter (E) followed by the reading of each question by the subject (S). Information relative to the question was given by E after the reading of each question. As E wished to keep himself open to answer any specific question or to correct any individual's misunderstandings in an informal tone and manner, there were no written instructions read to the Ss. A list of instructions to use with the Speech Ruler was developed in the pre-experimental pilot stages of this project. This list was memorized by E and followed carefully in instructing the S in the use of the Speech Ruler. In this manner E was able to give each S virtually the same instructions while maintaining an informal tone to the communications. An audio-tape record was kept of these instructions for the first Ss in the experimental series. An examination of these indicated that E did not vary significantly from S to S in giving the instructions. Such variations that existed were due primarily to the explanation of specific terms to individual Ss.

What follows in a transcript of E's explanations on the use of the Speech Ruler to a S. Notes in parentheses refer to deliberate movements carried out by E. The script has been edited with the deletion of certain verbal and non-verbal expressions to make it more print-readable.

A copy of the Speech Ruler may be found in Appendix D.

E: (Hands S a copy of the Speech Ruler; holds one copy for himself.) First, let's go over this Speech Ruler. I'll read a little of it to you, then you'll read the rest. (Reading opening paragraph.) "These questions will help you measure your own speech. Just as you can measure how tall you are with a ruler, you can measure how good a speaker you are with this Speech Ruler. Read each question carefully. Think about your speech. Then place a check mark on the Ruler where you measure yourself to be. For example, if I ask you how fast you can run, give me the answer by placing a mark like this on the Ruler." This is the Ruler (E points to example). You're going to tell me how fast or how slow you can run. If you're a slow runner you put a mark over here at this end of the Ruler (E points to first sector of Ruler line). If you're a fast runner place a mark at this end (E points to last sector of Ruler). Now, what if you are neither fast nor slow but just average? Where would you put your mark?

S: (Answers verbally or by pointing. If not correct or uncertain, E repeats and elaborates.)

E: Good. What if you're faster than average but not the fastest runner?

S: (Answers by pointing or indicating verbally. If S cannot answer then E repeats and elaborates. E might then ask for position of a slower-than-average-but-not-the-slowest runner.)

E: O.K. So now you tell me how fast a runner you are by placing a mark on the Ruler.

S: (Places a mark.)

E: (Comments on mark, e.g. "So you're about in the middle....") Good. Now to make sure you understand this I'm going to ask you to read the rest of the Ruler out loud to me once before marking it. I'll explain things to you as you go along. If you have any questions or if you don't understand a word ask me now because I'm not going to help you when you start judging for yourself. You're going to measure three things with the Speech Ruler. First, "What you said," these are your words, your ideas, not your voice. Second, "How you looked" and third, "How you sounded"--that's your voice. Now read the first one.

S: (Reads Ia of Speech Ruler.)

E: When you judge your speech I want you to ask yourself, "Did it make sense?", "Would people listening to it understand me?"

S: (Reads Ib.)

E: In a speech there are some ideas which are very important, which mustn't be left out. Suppose you forgot to say some things that are important you'd be at this end of the Ruler. (E points to first sector.) If you mentioned all the important ideas you'd be at that end (points to fifth sector).

S: (Reads Ic.)

E: Some people fill up their speech with only small ideas. Ideas that don't have much to do with the story. Other

people stick to the important ideas. They would be here on the Ruler (E points to fifth sector).

S: (Reads Id.)

E: A good speech is one which catches people's attention right at the beginning. A poor one starts off in a dull way.

S: (Reads Ie.)

E: Sometimes a speech ends...(Silence)...like that. People don't know it's over because it doesn't seem like it's finished. A good speech ending lets people know it's over.

Now before you read on, let's look back. You'll be making check marks. If you notice on this side (E points to the left section of Ruler) the Ruler is weak or poor. And on this side (E points to the right side) it is strong or good. I want you to be very careful when you are marking. Each question is different, so don't let your answer on one Ruler influence you on the next question. You could be very good on this one for example (E points to Ia.) but poor on that one (E points to Ib.) and average on this one (E points to Ic.). You could be different on each one. The reason I'm telling you this is that some people aren't careful. They may start off carefully marking a high on A, then high on B and then high on C and then they continue all the way down the page on one side of the Ruler without thinking. Others do the same thing down the left side of the page, or right down the middle. Now maybe you are good on each Ruler, or poor, or average. If you answer each question carefully

and end up with all your checkmarks on one side that's all right. What I don't want you to do is to mark them that way because you're not thinking. Remember you might be good on some things but poor on others. Now let's continue.

S: (Reads IIa.)

E: That has to do with: "Did you look like you wanted to speak?" If you look unhappy or bored, your audience won't be interested either in what you are saying.

S: (Reads IIb.)

E: Do you know what nervous means? It means looking scared or shakey.

S: (Reads IIc.)

E: Some people look at the floor, the ceiling, to the side but never at the people they are talking to. You should look at your audience. Here, that means me and the camera. Some people make a different kind of mistake on where they look. They seem to be looking at the audience but they are really looking through them. It's as if they were alone and talking to themselves (E speaks last sentence in the vacant manner described) like I'm doing right now.

S: (Reads IId.)

E: Some people talk like this: they don't move (E sits stiffly, arms pressed against side). Other people are relaxed; they move their hands or body to sometimes show what they are talking about. Try not to look like a statue of stone.

S: (Reads IIe.)

E: Do you know what "annoying" means? It means disturbing. Annoying movements are movements which have nothing to do with the speech. For example, some people pull their ear like this, (E tugs at ear), all during their speech, although they're not talking about ears. There are all kinds of small annoying movements which you have to avoid in a speech.

S: (Reads IIf.)

E: This has to do with a thing called "expression." Having an expression on your face means making your face show what you are talking about. If your speech is happy, you should look happy. If it's sad, you should look sad. Some people have no expression on their face. They keep the same face all the time, no matter what they say.

S: (Reads IIIa. and IIIb.)

E: Now these last two questions are just like two we had before (E points back to IIa.and b.) which asked, "Did you look interested?", "Did you look nervous?" Now the Ruler is asking, "Did you sound interested?", "Did you sound nervous?"

S: (Reads IIIc.)

E: This ruler and the next are a little different. It's twice as long as the others because you could make a mistake in two ways. You could be too loud, (E speaks loudly), or not loud enough, (E speaks softly). Where would you be if you spoke loud enough?

S: (Reads IIId.)

E: This question is like the last.

S: (Reads IIIe.)

E: Do you know what "pronounce" means? For example, if you mumble your words, (E mumbles this last phrase), it may not be understood by your audience. You should always try to speak clearly.

S: (Reads IIIf.)

E: We saw what an annoying movement was before. It was disturbing. An annoying sound is the same thing. It's a sound that doesn't belong in your speech like "eh," "ah," "(cough)" "(heavy sigh)."

S: (Reads IIIg.)

E: We saw what an expression on the face was. Expression in the voice is the same thing. Some people have no expression. Their voice sounds like this: (E Makes a long flat hum). Others make their voice go up and down, so that if they are talking about something happy, their voice sounds happy. If they are talking of something sad, their voice sounds sad.

The last two questions deal with your whole speech. How do you judge your whole speech?--not just your voice, or how you look, or what you said, but the whole speech.

S: (Reads IVa.)

E: Is that clear?

S: (Reads IVb.)

E: By this I mean: Suppose you are very satisfied with your speech. If it's the best one, you,_____ (S's name, e.g. Billy Gold) can give, mark it over here, (E points to sector 5). This doesn't mean it's the <u>best speech</u> in the

world. Maybe somebody else like Mrs. Johnston (School Principal), or Mr. Rattray (a teacher), can give a better speech. But if it's your best speech, mark it there. If, on the other hand, you know you can do better, that it's not your best speech, mark it over here, (E points to first sector). This doesn't mean it's the worst speech in the world. Some people might do a worse job of it. If it's neither you at your best nor you at your worst, mark it some place in the middle.

O.K. Are there any questions?

APPENDIX D

THE SPEECH RULER

Subjects in groups VTR-CHKL and CHKL used the Speech Ruler, a self-evaluative checklist, to evaluate their own speech delivery performances. A copy of the Speech Ruler is reproduced below. The subjects' copy was identical with the exception of being contained on two $8\frac{1}{2}$ x 14 inch sheets.

The Speech Ruler

These questions will help you measure your own speech. Just as you can measure how tall you are with a ruler, you can measure how good a speaker you are with this Speech Ruler.

Read each question carefully. Think about your speech. Then place a check mark on the Ruler where you measure yourself to be. For example, if I ask you how fast you can run, give me the answer by placing a mark like this on the Ruler.

very slowly :____:____:____:____:____: very quickly

Now let's measure your speech.

I. What was said

a) Did your speech make sense? Would everybody understand what you wanted to say?

No, it was not clear :____:____:____:____:____: Yes, it was very clear

b) Did you say everything that was important? Was anything important missing?

It was not complete :____:____:____:____:____: It was very complete

c) Did you say some things that were not necessary? Did you say too much?

There were too many small ideas :____:____:____:____:____: Everything I said was important

d) How was the beginning? Could it catch people's attention?

It was dull :____:____:____:____:____: It was very interesting

e) How was the ending? Would people know that the speech was finished?

It did not seem like an ending : ____ : ____ : ____ : ____ : ____ : It ended the speech very well

II. How you looked

a) Did you look interested in what you were saying? Did you look pleased to be speaking?

No : ____ : ____ : ____ : ____ : ____ : Yes

b) Did you look nervous?

I looked very nervous : ____ : ____ : ____ : ____ : ____ : I looked relaxed

c) Did you look at the audience? Did you look as if you were talking to the audience or to yourself?

I hardly looked at the audience : ____ : ____ : ____ : ____ : ____ : I always looked at the audience

d) Did you move your hands and body enough to show what you were saying?

No, I was very stiff : ____ : ____ : ____ : ____ : ____ : Yes, I moved enough

e) Did you move too much? Did you have any annoying movements?

Yes, my movements were annoying : ____ : ____ : ____ : ____ : ____ : No, I used just the right movements

f) Did your face show what you were talking about?

No, it never changed : ____ : ____ : ____ : ____ : ____ : Yes, it changed enough

III. How you sounded

a) Did you sound interested in what you were saying? Did you sound pleased to be speaking?

No : ____ : ____ : ____ : ____ : ____ : Yes

b) Did your voice sound nervous?

Yes, very nervous : ____ : ____ : ____ : ____ : ____ : No, not at all

c) How loud was your voice?

Not loud enough : ____ : ____ : ____ : ____ : ____ : ____ : ____ : ____ : ____ : Too loud

just right

d) How fast did you speak?

Not fast enough : ____ : ____ : ____ : ____ : ____ : ____ : ____ : ____ : ____ : Too fast

just right

e) Was your voice clear? Were your words well pronounced?

No, not very clear : ____ : ____ : ____ : ____ : ____ : Yes, very clear

f) Did you make any unnecessary or annoying sounds?

Yes, too many : ____ : ____ : ____ : ____ : ____ : No, none at all

g) Did your voice have expression? Did it change enough to match what you said?

No, it stayed the same all the time : ____ : ____ : ____ : ____ : ____ : Yes, it had a lot of expression

IV. General

a) On the whole, how do you judge your speech?

Very poor : ____ : ____ : ____ : ____ : ____ : Very good

b) How much better do you think you could give this speech?

I could do much better than this : ____ : ____ : ____ : ____ : ____ : This was my best

APPENDIX E

SPEECH CHECKLIST

The Speech Checklist was the evaluative form used by the three judges in the experiment. Each subjects' five videotaped speeches were evaluated by each judge on this form. The judges' form was contained on one $8\frac{1}{2}$ X 14 inch sheet.

IRPQ/Boyd, Jones/July 21, 70

____ ____ ____ ____ ____

Speech Checklist

I. CONTENT

a) Clarity:	Made little sense	:____:____:____:____:____:	Made sense
b) Coverage:	Not complete	:____:____:____:____:____:	Very complete
c) Verbosity:	Too much irrelevance	:____:____:____:____:____:	Everything said was important
d) Beginning:	Dull	:____:____:____:____:____:	Interesting
e) Ending:	Poor	:____:____:____:____:____:	Good

II. DELIVERY: VISUAL

a) Appearance:	Did not look interested	:____:____:____:____:____:	Looked interested
b) Nervousness:	Yes	:____:____:____:____:____:	No
c) Audience-contact:	Poor	:____:____:____:____:____:	Good
d) Body Movement:	Stiff	:____:____:____:____:____:	Expressive
e) Annoying Movements:	Many	:____:____:____:____:____:	None
f) Facial Expression:	None	:____:____:____:____:____:	Enough

III. DELIVERY: VOCAL

a) Tone	Did not sound interested	:____:____:____:____:____:	Sounded interested
b) Nervousness:	Yes	:____:____:____:____:____:	No
c) Loudness:	Not loud enough	:____:____:____:____:____:____:____:____:____: (just right)	Too loud
d) Speed:	Not fast enough	:____:____:____:____:____:____:____:____:____: (just right)	Too fast

e) Pronounciation: Poor :____:____:____:____:____: Good

f) Annoying Sounds: Many :____:____:____:____:____: None

g) Vocal Expression: None :____:____:____:____:____: Enough

IV. GENERAL

a) Overall: Very poor :____:____:____:____:____: Very Good

b) Improvement: Much needed :____:____:____:____:____: Little needed

APPENDIX F

ORAL EVALUATION ANALYSIS FORM

All subjects gave oral evaluations of their speeches. These were recorded on audio cassette and subjected at a later date to a simple content analysis measuring the frequency of types of statements. A copy of the form used to direct this analysis is reproduced below.

Oral Evaluation Analysis

____ ____ ____ ____ ____

I. Reaction to Self

___1. Looked like self_____

___2. Did not look like self_____

___3. Sounded like self_____

___4. Did not sound like self_____

II. Evaluation of Speech

SPEECH NO. CATEGORY	1	2	3	4	5	TOTAL
1. Content						
2. Visual						
3. Vocal						
4. General						
TOTAL						

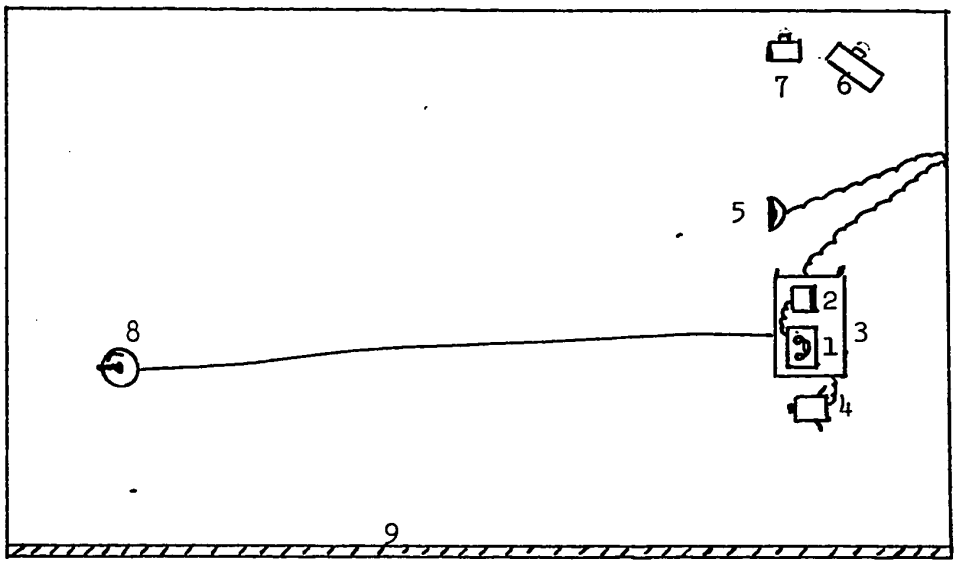
III. Overall Evaluation

1. My best speech

2. My worst speech

APPENDIX G

VIDEO RECORDING AND PLAYBACK SET-UP



1. Videotape recorder
2. Monitor
3. Mobile table
4. Camera
5. Floodlight
6. Investigator's desk
7. Subject's desk
8. Microphone
9. Stage curtain

APPENDIX H

MEANS OF JUDGES' EVALUATIONS, SELF-EVALUATIONS AND DISCREPANCY SCORES OVER FIVE TRIALS FOR THE FOUR TREATMENT GROUPS

The three judges' total scores for each subject were averaged yielding the "Judges" data of Table 14. The subjects' self-evaluations as recorded in the Speech Ruler yielded the "Self" data. Subtracting the self-evaluation scores from the judges' mean scores yielded the "Discrepancy" data.

TABLE 14.--Means of judges' evaluations, self-evaluations and discrepancy scores over five trials for the four treatment groups

Group	Subject	Judges					Self					Discrepancy				
		Speech No.					Speech No.					Speech No.				
		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
CHKL	A061	50	48	51	50	48	93	97	97	95	96	43	49	46	45	48
	A147	53	50	50	53	48	62	77	78	76	81	9	27	28	23	33
	A160	61	48	59	54	53	63	50	77	77	74	2	2	18	23	21
	A034	64	56	61	68	60	77	82	77	82	87	13	26	16	14	27
	A152	50	50	45	42	46	61	81	88	86	90	11	31	43	46	44
	A008	50	58	54	61	60	55	74	74	86	93	5	16	20	25	33
	A163	59	54	59	59	60	68	66	75	80	94	9	12	16	21	34
	A109	51	46	50	47	53	93	100	98	98	96	42	54	48	51	43
	A030	45	48	51	54	50	68	61	55	68	79	23	13	4	14	29
	A029	47	47	58	54	44	77	83	86	87	91	30	36	28	33	47
	A054	57	63	54	61	48	77	87	86	83	82	20	24	32	22	34
	A009	60	68	56	68	57	82	84	97	99	99	22	16	41	31	42
	A117	69	63	60	65	66	38	52	83	64	84	-31	-11	23	- 1	18
	A012	56	64	51	53	62	32	56	72	54	78	-24	- 8	21	1	16
	A092	51	59	60	60	59	84	86	92	91	94	33	27	32	31	35
	A167	62	70	74	61	61	84	97	100	97	100	22	27	26	36	39

TABLE 14.--Continued

Group	Subject	Judges					Self					Discrepancy				
		Speech No.					Speech No.					Speech No.				
		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	A057	40	50	42	47	44	59	63	59	57	57	19	13	17	10	13
	A105	46	51	51	47	49	65	73	49	63	76	19	22	- 2	16	27
	A085	50	53	48	52	53	80	81	84	85	88	30	28	36	33	35
	A089	40	44	44	42	42	87	80	86	91	84	47	36	42	49	42
VTR-CHKL	B166	56	56	64	65	61	53	72	88	86	89	- 3	16	24	21	28
	B001	69	70	65	65	68	69	81	82	89	93	0	9	17	24	25
	B006	59	51	63	69	67	67	81	88	96	97	8	30	25	27	30
	B013	49	51	50	53	46	50	61	59	81	61	1	10	9	28	15
	B143	63	57	54	60	57	80	88	86	76	81	17	31	32	16	24
	B052	53	56	55	63	58	67	65	81	87	88	14	9	26	24	30
	B055	62	60	67	62	63	74	65	67	73	54	12	5	0	11	- 9
	B136	52	51	56	57	56	56	59	87	63	69	4	8	31	6	13
	B059	52	56	57	63	58	65	74	88	91	91	13	18	31	28	33
	B005	62	54	65	64	57	56	66	70	78	77	- 6	12	5	14	20
	B035	60	72	74	73	78	51	73	86	88	99	- 9	1	12	15	21
	B041	49	41	46	50	48	48	55	68	77	52	- 1	14	22	27	4

TABLE 14.--Continued

Group	Subject	Judges					Self					Discrepancy				
		Speech No.					Speech No.					Speech No.				
		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	B142	68	61	68	72	75	80	60	86	88	86	12	- 1	18	16	11
	B021	55	62	66	67	66	55	74	74	64	78	0	12	8	- 3	12
	B133	57	53	49	59	53	60	66	80	84	77	3	13	31	25	24
	B155	80	80	78	75	81	82	89	96	96	100	2	9	18	21	19
	B166	53	55	71	65	80	48	82	63	82	90	- 5	27	- 8	17	10
	B047	60	75	70	77	76	39	50	52	58	69	-21	-25	-18	-19	- 7
	B135	53	63	62	66	64	70	71	70	91	92	23	8	8	25	28
	B172	66	60	67	61	64	58	80	90	57	100	- 8	20	23	- 4	36
	B157	64	72	68	61	63	72	96	88	82	100	8	24	20	21	37
VTR	C113	33	40	47	39	52										
	C064	42	44	46	37	47										
	C096	22	58	62	61	55										
	C036	35	42	40	43	49										
	C088	35	36	36	44	37										
	C139	62	52	55	56	58										
	C121	60	60	55	59	57										

TABLE 14.--Continued

Group	Subject	Judges					Self					Discrepancy				
		Speech No.					Speech No.					Speech No.				
		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C115	58	69	47	61	57										
	C169	44	42	46	45	37										
	C156	67	69	74	72	68										
	C164	58	62	56	59	58										
	C081	66	62	62	63	66										
	C173	57	56	55	59	57										
	C108	51	52	52	50	53										
CNTL	D168	52	53	46	51	54										
	D040	44	47	47	49	58										
	D124	61	62	56	57	58										
	D101	39	49	47	47	46										
	D048	59	59	68	69	71										
	D134	61	57	58	52	59										
	D004	53	47	46	48	51										
	D145	51	50	48	46	51										
	D060	38	39	37	44	38										
	D086	29	45	49	37	45										

TABLE 14.--Continued

Group	Subject	Judges					Self					Discrepancy				
		Speech No.					Speech No.					Speech No.				
		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	D075	50	43	56	55	52										
	D046	67	68	63	59	-[a]										
	D111	66	72	71	57	-[a]										
	D126	43	46	50	47	50										
	D124	49	55	47	54	56										

[a]No judges' evaluations for subjects D046 and D111

APPENDIX I

SEX, I.Q. AND PREVIOUS SPEECH TRAINING FACTORS

Although the hypotheses governing this research did not distinguish between types of subjects, data was available pertaining to the sex, intelligence test scores, and previous speech delivery training of the subjects. Since these data relate to some of the recommendations for future research and since curiosity prompts one to wonder what influence these elements may have had on the results reported in the main body of this report, these data were subjected to analyses and are reported here.

TABLE 15.--Means of judges' evaluations of male and female subjects

Speech Trial	Treatment Group							
	CHKL		VTR-CHKL		VTR		CNTL	
	Sex		Sex		Sex		Sex	
	M (n=10)	F (n=10)	M (n=10)	F (n=11)	M (n=5)	F (n=9)	M (n=7)	F (n=8)
1	49.0	57.1	55.8	62.2	39.2	54.9	47.2	55.0
2	50.5	58.5	56.7	62.6	46.4	56.9	48.9	57.6
3	51.0	56.8	58.2	66.1	47.8	54.9	50.1	55.4
4	52.2	57.6	62.0	67.1	48.2	56.3	48.4	55.0
5	49.4	56.9	60.1	64.9	49.2	56.1	51.4	53.6

TABLE 16.--Analyses of variance of judges' evaluations within each treatment group: trial X sex interaction

Group	Source of variation	Degrees of freedom	Sums of squares	Mean square	F	
CHKL	Trial	4	52.90	13.22	< 1	
	Sex	1	1211.04	1211.04	26.08	p<.01
	Trial X sex	4	32.26	8.06	< 1	
	Within	90	4178.80	46.43		
	Total	99	5475.00			
VTR-CHKL	Trial	4	440.68	110.17	1.45	n.s.
	Sex	1	1061.82	1061.82	14.02	p<.01
	Trial X sex	4	52.59	13.14	< 1	
	Within	95	7173.76	75.71		
	Total	104	8728.85			
VTR	Trial	4	179.90	44.97	< 1	
	Sex	1	1500.44	1500.44	14.61	p<.01
	Trial X sex	4	172.05	43.01	< 1	
	Within	60	6161.96	102.69		
	Total	69	8014.35			
CNTL	Trial	4	52.99	13.24	< 1	
	Sex	1	788.20	788.20	10.31	p<.01
	Trial X sex	4	65.90	16.47	< 1	
	Within	63	4814.04	76.41		
	Total	72	5721.13			

TABLE 17.--Means of self-evaluations of male and female subjects

Trial	Treatment Group			
	CHKL		VTR-CHKL	
	Sex		Sex	
	M (n=10)	F (n=10)	M (n=10)	F (n=11)
1	74.8	65.7	61.5	62.2
2	79.1	73.9	68.1	75.1
3	80.2	81.1	78.4	78.6
4	81.8	80.1	80.3	80.3
5	84.7	87.6	78.6	87.0

TABLE 18.--Analyses of variance of self-evaluations within treatment groups: trial X sex interaction

Group	Source of variation	Degrees of freedom	Sums of squares	Mean square	F	
CHKL	Trial	4	2001.70	500.42	2.55	$p<.05$
	Sex	1	148.84	148.84	< 1	
	Trial X sex	4	1267.46	316.86	1.61	n.s.
	Within	90	17653.00	196.14		
	Total	99	21071.00			
VTR-CHKL	Trial	4	6015.67	1503.91	9.49	$p<.01$
	Sex	1	287.10	287.10	1.81	n.s.
	Trial X sex	4	348.63	87.15	< 1	
	Within	95	15049.23	158.41		
	Total	104	21700.63			

TABLE 19.--Means of discrepancy scores of male and female subjects

Trial	Treatment Group			
	CHKL		VTR-CHKL	
	Sex		Sex	
	M (n=10)	F (n=10)	M (n=10)	F (n=11)
1	25.8	8.7	6.3	0.1
2	28.6	15.4	11.4	12.3
3	29.2	24.3	20.2	12.0
4	29.8	22.5	18.3	14.2
5	35.3	30.7	18.5	19.0

TABLE 20.--Analyses of variance of discrepancy scores within treatment groups: trial X sex interaction

Group	Source of variation	Degrees of freedom	Sums of squares	Mean square	F	
CHKL	Trial	4	2748.86	687.21	3.19	p<.05
	Sex	1	2218.41	2218.41	10.30	p<.01
	Trial X sex	4	607.14	151.78	< 1	
	Within	90	19373.50	215.26		
	Total	99	24947.91			
VTR-CHKL	Trial	4	3306.43	826.60	5.52	p<.01
	Sex	1	270.33	270.33	1.80	n.s.
	Trial X sex	4	384.04	96.01	< 1	
	Within	95	14215.26	149.63		
	Total	104	18176.06			

TABLE 21.--Means of judges' evaluations of subjects above and below the median I.Q.

Speech Trial	Treatment Group							
	CHKL		VTR-CHKL		VTR		CNTL	
	I.Q.		I.Q.		I.Q.		I.Q.	
	Below (n=11)	Above (n=9)	Below (n=9)	Above (n=10)	Below (n=7)	Above (n=7)	Below (n=7)	Above (n=8)
1	47.7	58.4	56.7	61.1	39.3	59.2	47.7	53.5
2	51.9	57.1	58.0	61.8	46.8	59.4	50.7	54.6
3	51.2	56.6	60.4	64.0	48.3	56.4	52.2	52.8
4	52.5	57.3	61.6	66.6	47.5	59.2	48.2	54.2
5	50.2	56.1	60.0	67.4	50.0	57.3	48.5	56.8

TABLE 22.--Analyses of variance of judges' evaluations within each treatment group: trial X I.Q. interaction

Group	Source of variation	Degrees of freedom	Sums of squares	Mean square	F	
CHKL	Trial	4	52.90	13.22	< 1	
	I.Q.	1	1024.00	1024.00	28.07	p<.01
	Trial X I.Q.	4	1114.30	278.57	7.63	p<.01
	Within	90	3283.80	36.48		
	Total	99	5475.00			
VTR-CHKL	Trial	4	419.19	104.79	1.10	n.s.
	I.Q.	1	79.59	79.59	< 1	
	Trial X I.Q.	4	48.89	12.22	< 1	
	Within	85	8045.06	94.64		
	Total	94	8592.73			
VTR	Trial	4	179.90	44.97	< 1	
	I.Q.	1	2496.06	2496.06	30.05	p<.01
	Trial X I.Q.	4	355.22	88.80	1.06	n.s.
	Within	60	4983.17	83.05		
	Total	69	8014.35			
CNTL	Trial	4	52.99	13.25	< 1	
	I.Q.	1	423.67	423.67	5.20	p<.01
	Trial X I.Q.	4	118.13	29.53	< 1	
	Within	63	5126.34	81.37		
	Total	72	5721.13			

TABLE 23.--Means of self-evaluation of subjects above and below the median I.Q.

Speech Trial	Treatment Group			
	CHKL		VTR-CHKL	
	I.Q.		I.Q.	
	Below (n=11)	Above (n=9)	Below (n=9)	Above (n=10)
1	74.5	66.0	61.4	63.2
2	78.8	74.2	72.1	70.9
3	76.9	84.4	81.3	75.1
4	80.9	81.0	82.4	80.5
5	84.0	88.3	86.4	79.0

TABLE 24.--Analyses of variance of self-evaluations within each group: trial X I.Q. interaction

Group	Source of variation	Degrees of freedom	Sums of squares	Mean square	F	
CHKL	Trial	4	2001.70	500.43	2.58	p<.05
	I.Q.	1	1.44	1.44	< 1	
	Trial X I.Q.	4	1645.86	411.47	2.12	n.s.
	Within	90	17422.00	193.57		
	Total	99	21071.00			
VTR-CHKL	Trial	4	4913.17	1228.29	7.52	p<.01
	I.Q.	1	191.44	191.44	1.17	n.s.
	Trial X I.Q.	4	248.99	62.24	< 1	
	Within	85	13880.13	163.29		
	Total	94	19233.73			

TABLE 25.--Means of discrepancy scores of subjects above and below the median I.Q.

Speech Trial	Treatment Group			
	CHKL		VTR-CHKL	
	I.Q.		I.Q.	
	Below (n=11)	Above (n=9)	Below (n=9)	Above (n=10)
1	26.9	7.6	5.1	3.2
2	26.9	17.1	12.0	10.8
3	25.7	27.8	19.7	12.4
4	28.4	23.9	19.1	15.7
5	33.8	32.2	25.1	13.6

TABLE 26.--Analyses of variance of discrepancy scores within each group: trial X I.Q. interaction

Group	Source of variation	Degrees of freedom	Sums of squares	Mean square	F	
CHKL	Trial	4	2748.86	687.21	3.13	p<.05
	I.Q.	1	1095.61	1095.61	5.00	p<.05
	Trial X I.Q.	4	1383.15	345.78	1.57	n.s.
	Within	90	19720.29	219.11		
	Total	99	24947.96			
VTR-CHKL	Trial	4	2457.62	614.40	4.09	p<.01
	I.Q.	1	562.67	562.67	3.75	n.s.
	Trial X I.Q.	4	327.78	81.94	< 1	
	Within	85	12751.77	150.02		
	Total	94	16099.84			

TABLE 27.--Means of judges' evaluations of subjects with and without previous speech training

Speech Trial	Treatment Group							
	CHKL		VTR-CHKL		VTR		CNTL	
	Speech Training		Speech Training		Speech Training		Speech Training	
	No (n=11)	Yes (n=9)	No (n=10)	Yes (n=11)	No (n=9)	Yes (n=5)	No (n=8)	Yes (n=7)
1	47.9	59.3	57.3	60.8	44.3	58.2	45.2	57.1
2	51.5	58.1	58.5	61.0	51.0	57.0	48.5	57.7
3	51.1	57.2	61.3	63.8	49.6	57.2	50.4	55.1
4	52.3	58.1	63.2	65.0	50.8	58.2	48.0	55.4
5	50.0	57.0	61.7	65.6	52.5	55.6	48.5	58.1

TABLE 28.--Analyses of variance of judges' evaluations within each treatment group: trial X training interaction

Group	Source of variation	Degrees of freedom	Sums of squares	Mean square	F	
CHKL	Trial	4	52.90	13.22	< 1	
	Training	1	1345.69	1354.69	21.73	p<.01
	Trial X training	4	145.60	36.40	< 1	
	Within	90	3930.81	43.67		
	Total	99	5475.00			
VTR-CHKL	Trial	4	440.68	110.17	1.29	n.s.
	Training	1	213.41	213.41	2.51	n.s.
	Trial X training	4	15.53	3.88	< 1	
	Within	95	8060.23	84.84		
	Total	104	8728.85			
VTR	Trial	4	179.90	44.97	< 1	
	Training	1	921.79	921.79	8.24	p<.01
	Trial X training	4	201.27	50.31	< 1	
	Within	60	6711.39	111.85		
	Total	69	8014.35			
CNTL	Trial	4	52.99	13.24	< 1	
	Training	1	1327.32	1327.32	19.74	p<.01
	Trial X training	4	106.01	26.50	< 1	
	Within	63	4234.81	67.21		
	Total	72	5721.13			

TABLE 29.--Means of self-evaluations of subjects with and without previous speech training

Speech Trial	Treatment Group			
	CHKL		VTR-CHKL	
	Training		Training	
	No (n=11)	Yes (n=9)	No (n=10)	Yes (n=11)
1	76.1	63.0	57.7	65.7
2	80.4	71.6	65.9	77.1
3	78.7	83.0	75.8	81.0
4	82.1	79.4	76.0	84.2
5	85.0	87.4	78.0	87.5

TABLE 30.--Analyses of varience of subjects' self-evaluations within each group: trial X training interaction

Group	Source of variation	Degrees of freedom	Sums of squares	Mean square	F	
CHKL	Trial	4	2001.70	500.42	2.67	p<.05
	Training	1	323.64	323.64	1.72	n.s.
	Trial X training	4	1880.09	470.02	2.50	p<.05
	Within	90	16865.57	187.39		
	Total	99	21071.00			
VTR-CHKL	Trial	4	6015.67	1503.92	10.42	p<.01
	Training	1	1875.91	1875.91	13.00	p<.01
	Trial X training	4	105.70	26.42	< 1	
	Within	95	13703.35	144.24		
	Total	104	21700.63			

TABLE 31.--Means of discrepancy scores of subjects with and without previous speech training

Speech Trial	Treatment Group			
	CHKL		VTR-CHKL	
	Training		Training	
	No (n=11)	Yes (n=9)	No (n=10)	Yes (n=11)
1	28.3	3.7	1.0	4.9
2	28.9	13.5	7.4	16.0
3	27.5	25.7	14.5	17.1
4	29.5	21.5	12.8	19.2
5	35.0	30.4	16.3	21.9

TABLE 32.--Analyses of variance of discrepancy scores within each group: trial X training interaction

Group	Source of variation	Degrees of freedom	Sums of squares	Mean square	F	
CHKL	Trial	4	2748.86	687.21	3.52	p<.05
	Training	1	2974.98	2974.98	15.26	p<.01
	Trial X training	4	1678.81	419.70	2.15	n.s.
	Within	90	17545.26	194.94		
	Total	99	24947.91			
VTR-CHKL	Trial	4	3306.43	826.60	5.61	p<.01
	Training	1	779.22	779.22	5.29	p<.05
	Trial X training	4	110.16	27.54	< 1	
	Within	95	13980.25	144.16		
	Total	104	18176.06			