Language Studies of Mentally Retarded Children

A Report of the Parsons Project in Language and Communication of Mentally Retarded Children

Richard L. Schiefelbusch, Ph.D., Project Director

Howard V. Bair, M.D., Project Co-Director

Joseph E. Spradlin, Ph.D., Field Director

Gerald M. Siegel, Ph.D.

Frances D. Horowitz, Ph.D.

Jerome P. Harkins, M.A.

Ross H. Copeland, M.S.

The Journal of Speech and Hearing Disorders
Monograph Supplement Number 10

January, 1963

The American Speech and Hearing Association

The Journal of Speech and Hearing Disorders Monograph Supplement Number 10

Language Studies of Mentally Retarded Children

Appendixes

Preface

The Parsons Project in Language and Communication of Mentally Retarded Children[1] was initially developed by a research planning committee composed of representatives from the Menninger Foundation, Parsons State Hospital and Training Center, and the Bureau of Child Research of the University of Kansas.

The purpose of the project, as stated, was to develop a language and communication research program for a selected group of educable mentally retarded children in an institutional setting. The project planners chose to focus upon the effects that environmental, especially social, factors may have in shaping a retarded child's language behavior. Five subgoals were included in the initial project statement:

1. The formulation of a set of experimental constructs relative to language as a feature of social adequacy.
2. Diagnostic assessments which include descriptions of the negative and maladaptive behavior patterns affecting language and communication.
3. The development of a battery of language and communication tests.
4. The development of specific clinical techniques for improving the verbal behavior of mentally retarded children.
5. The development of a team program for improving the environmental milieu of the institutional child for purposes of stimulating verbal development.

The research staff subsequently undertook to translate these aims into empirical studies and to extend them or to reinterpret them as a feature of the ongoing research effort.

The first three years of the project are reflected in this monograph. Subsequent research reports will be needed to show the scope and diversity of work in the Parsons Research setting during the second three-year period. It is apparent now that the processes of research have led the staff to make increasingly more realistic assessments of variables and thus to formulate increasingly perceptive research questions.

The current staff of the project includes the following: Field Director, Joseph E. Spradlin; Research Associates: Ross Copeland, John deJung, Fred Girardeau, John Hollis, and Frances Horowitz; Research Assistants: Clara Bauer, Donald Carter, Donald Dickerson, Eric Errickson, Mary Schickel, and Ruth Staten; Research Fellow: Bill Locke. Seymour Rosenberg, first Field Director, and Dorolyn Ezell and Gerald Siegel, Research Associates, were formerly associated with the project.

No large research undertaking can succeed without the cooperation of many people. This is particularly true in early planning phases. Therefore, we would like to acknowledge the interest and aid of the people who were part of the original advisory committee: Richard Bartman, Margaret Byrne, John W. Fair, Bernard Foster, Cotter Hirschberg, Lloyd Lockwood, Paul Pruyser, John Segerson, and Erik Wright. In addition, the encouragement of George W. Jackson, former Superintendent of Institutions, was most helpful.

[1]Supported by funds from the National Institutes of Mental Health (OM-111).

The relative ease with which the research program was fitted into the already existing structure of the Parsons State Hospital and Training Center was due to the full cooperation and acceptance on the part of the staff members of the Center. In this respect, we would particularly like to acknowledge the help of Chester Gorton, Henry Leland, and Lloyd Lockwood and their staffs.

Harris Winitz has lent important assistance throughout and has played an important critical role in the development of this monograph. We wish also to acknowledge the contribution of Dr. Seymour Rosenberg, field director of the Project from June, 1958, to September, 1959. His able leadership and abilities as a creative scientist were critically important during the early phases of the project formulation. Finally, the secretarial and detail efforts of Ruth Staten and her assistants, Mary Schickel and Clara Bauer, have been appreciated.

Richard L. Schiefelbusch, Ph.D.,
Project Director
Howard V. Bair, M.D.,
Project Co-Director

Introduction

RICHARD L. SCHIEFELBUSCH

Importance of the Problem

Communication deficiencies are a recognized part of the behavior of mentally retarded children. Speech and language surveys of institutionalized retardates report that these children are deficient in vocabulary, sentence structure, conceptual and abstract language skills, voice quality, and the articulation of speech sounds (*2*). A survey of speech deficits by Spradlin (*12*) leads to the following generalizations:

1. From 57 to 72% of institutionalized mental defectives have speech defects.
2. Approximately 72 to 82% of severely retarded children in parent-sponsored day schools have speech defects.
3. Eight to 26% of the children in special classes of the public schools have speech defects.
4. Articulation and voice problems comprise the largest percentage of speech problems among mentally retarded children.

The reviews of Matthews (*2*) and Spradlin (*12*) emphasize the extent of speech defects among mentally defective persons. However, non-speech communication behavior, such as use of gestures and responses to speech and gestures, very likely play an important role in the communication of mentally retarded children. These aspects of communication have received far less attention. In fact, Spradlin was unable to find a single comprehensive study of these non-speech aspects of communication.

Attempts at remediation of language functioning also appear to have been hesitant and limited in focus. The lack of widespread professional enthusiasm for this work may be due in part to the assumption among investigators that the behavioral defects of the mentally defective person are due solely to innate organic deviations. Such assumptions do not serve as a stimulating foundation for either behavioral research or rehabilitative programs.

In spite of the physiological emphasis which is usually built into diagnostic processes, rehabilitative procedures almost invariably refer to environmental modifications aimed at behavioral change. A systematic approach from an environmental point of view would initially involve the identification of relevant variables associated with retarded functioning and with methods for effecting behavior change. From this point of view communication problems can be investigated as problems of verbal learning, and educational or clinical procedures can be considered learning programs within an interpersonal setting. If measures of progress are determined

Richard L. Schiefelbusch (Ph.D., Northwestern University, 1951) is Professor of Speech and Director of the Bureau of Child Research, University of Kansas. He is also Director of the Parsons Project. Co-director of the Parsons Project is Howard V. Bair (M.D., University of Kansas, 1943), Superintendent and Medical Director of the Parsons State Hospital and Training Center.

from behavioral changes the child makes within the interpersonal setting, it is possible to view a wider range of events as pertinent to the remedial process. 'Relevant behavior change' can then be substituted for 'normal speech' as a goal in training.

The apparent apathy which speech clinicians have shown for work with retardates, then, may be due as much to their interpretation of objectives and clinical procedures as to the behaviors of the retardate. It is possible that the application of systematic procedures within a behavioral or an interpersonal framework may eventually provide new clinical perspectives and techniques.

The research reported here represents initial approaches to various systematic programs of research within an environmental framework. The Parsons Project should be viewed as a basic research setting in which a series of studies will hopefully contribute information useful to the complex task of improving communication behavior. In the remaining portion of this introduction, attention will be focused on early developmental stages of the project and on guiding principles that characterize the research; finally a brief introduction to the specific sections is included to lay the groundwork for more extensive research reports.

Conceptual Sources of the Parsons Language Research

The staff of the Parsons Project has been drawn from diverse professional backgrounds, disciplines, and training programs. For this reason an initial task was the acquisition of a common body of basic information in the fields under inquiry and a common vocabulary with which to discuss this information. A series of seminars and discussions was initiated to serve a number of vital functions. Even more important than the transmission of information, perhaps, they provided the circumstance in which staff members exposed their differences in background and orientation and learned to translate each other's professional jargon into more personally familiar terms. The professional 'in group' language of the staff was ultimately valuable in the important planning phases of the empirical studies (*8*), (*9*).

A second task of the research staff was to develop a rationale for the empirical studies to be initiated. The rationale was drawn primarily from two sources: modern learning theory and social or interpersonal psychology (*1*, *3*, *4*, *5*, *6*, 7, *10*, *11*). Within learning theory language is considered to be subject to the control both of cues and reinforcement. The cue or discriminative stimulus refers to the situational conditions under which a response will be followed by reinforcement. Reinforcement is any event which changes the frequency of the response which it follows. Section One is essentially a study of individual differences in language response of mentally retarded children to a set of standardized cues. Section Three presents studies of the effect of certain reinforcements or consequences on language behavior.

Within the social or interpersonal conceptualization, verbal behavior is considered a response sequence between two persons. Failures in verbal learning by the child are evaluated in terms of the behavior of both the child and the other person. An interpersonal approach sug-

gests that adults, no matter how well instructed, may be affected by social processes which may modify their behavior. This modification may further mediate against appropriate verbal learning by the child. The studies reported in Section Two are a reflection of this interpersonal point of view.

While much of the research of the staff has been derived from these two frames of reference, studies have and will continue to be planned from other orientations.

Despite the diversity in specific research methods and interests, reflected in this monograph, certain fundamental assumptions have guided much of the research and may be summarized as follows:

(a) The study of language and communication will be facilitated if the terms are defined so that the events to which they refer can be observed, classified, and measured, that is, defined operationally.

(b) Language and communication behavior is determined by other events which can be objectively described, classified, and, in many instances, manipulated.

(c) Language behavior is learned and as such is subject to the same principles as other behavior.

Description of the Setting and the Research Population

The research reported in this monograph was conducted at the Parsons State Hospital and Training Center. The Center was organized for resident treatment in 1953 to meet the needs of educable, ambulatory mental defectives between the ages of six and 21. The goals of the institution include the return of a large percentage of the children to the community. Consequently, residents are, where feasible, provided a program of therapy and/or training. Children live in self-contained cottages with recreational and cooking facilities though main meals are taken in common dining areas.

The present population is 670 moderately retarded and emotionally disturbed boys and girls. The Hospital's staff numbers 405 employees. The following departments and disciplines are represented: Psychiatry, Medicine, Pediatrics, Psychology, Social Service, and Adjunctive Therapies. The Department of Adjunctive Therapies consists of special education, occupational therapy, music therapy, recreational therapy, bibliotherapy, religious education, vocational counseling, pre-vocational training, and speech pathology and audiology. The nursing service is composed of 180 trained psychiatric aides and 18 registered nurses.

At the time the Parsons Project was begun in 1958 the research area included a suite of 11 rooms located primarily in one wing of the hospital building. The rooms included individual offices, a large group office including a small reprint and book library, a conference room, experimental rooms, observation rooms, a sound control booth, a sound-proof audiometric room, and storage and workshop space. Since that time an experimental cottage, an operant conditioning suite, and two additional offices have been added.

Summaries of Areas to be Reported

The studies reported in this monograph were conducted between August, 1958, and September, 1961, and relate to three broad areas of ongoing language research. These areas are concerned with

language measurement of mental defectives (Section One); dyadic and small group studies where language output was investigated as a function of group composition (Section Two); and the relationship of reinforcing variables to language performance and learning (Section Three). More recently, but not reported in the present volume, studies have been added to the Parsons Research Project in the areas of comparative psychology, discrimination learning, and social interaction in cottage living areas.

Section One describes a procedure for sampling language behavior called The Parsons Language Sample (PLS). The procedure is based on Skinner's classification system: mand, tact, and verbal behavior under the control of verbal stimuli.

The total test consists of 123 items divided into seven subtests. Three of the subtests—*tact*, *echoic*, and *intraverbal*—sample the child's vocal or speech behavior. Three others—*echoic gesture*, *comprehension*, and *intraverbal-gesture*—measure non-vocal communication. The seventh subtest, *mand*, measures either vocal or non-vocal behavior.

Empirical evaluations of the PLS have included tests for examiner equivalence, reliability (including split half and test-retest coefficients), intersubtest correlations, and correlations of the PLS with IQs of the Weschler scales and with ratings of the children's behavior in nontest situations. The PLS has been used extensively in the Parsons research setting to select and group children for research studies and to measure change during treatment studies.

Section Two reports the results of four 'assembly' studies conducted at Parsons from 1959 through the spring of 1961. These studies were based on the hypothesis that adults may respond to poor verbalizers in such a way as to perpetuate a low level of verbal performance. In the first study by Spradlin and Rosenberg, the adult was given instructions to 'interview' the child. It was hypothesized that children of low verbal levels would condition adults to ask questions that require only a binary or two-choice answer, for example, 'yes' or 'no,' and that the questions asked of children of higher verbal levels would be more 'open ended.' Thus the primary concern was with the ways in which verbal levels of children may selectively influence the linguistic patterns of adults.

In a second study Siegel assembled adults with high verbal and low verbal children in a series of permissive 'play therapy-like' sessions extending over approximately 12 weeks.

In a third study by Siegel and Harkins adults participated with high and low children in an unstructured and a structured tutorial period. Again focus was on the effects of the children on the adults.

The last study of Siegel pertained to two modes of adult behavior: a permissive 'clinical approach' and an interview approach. The author hypothesized that the adults using the former method would elicit more child responses than they would under interview conditions. Different sets of pre-experimental instructions were given to the two groups of adults. Effects of the two conditions were determined by measures of the children's responses. In addition, adult verbal behavior was measured in relationship to the verbal levels of the children.

The third section describes two studies which investigate the effects of modes of reinforcement on verbal be-

havior of mentally retarded children. The first, 'Partial and Continuous Reinforcement of Verbal Responses Using Candy, Vocal, and Smiling Reinforcers among Retardates' by Horowitz, was concerned with the variables which affect the learning of vocal responses. Retardates were required to learn one of three vocal responses. A correct response was followed by candy, a vocal commendation, a smile, a vocal commendation plus candy, or a vocal commendation plus a smile. Further, one half of the subjects received continuous reinforcement and one half received a partial reinforcement. The results were studied with reference to the effects of these reinforcing stimuli and the effect of partial reinforcement on the learning of retardates. In general, the study attempted to determine the relationship of the specified reinforcing variables to the learning of vocal responses.

The second study described in the third section is 'The Effects of Free-Field Feedback Modification of Verbal Behavior' by Copeland. It describes the author's 'immediate feedback' procedure of feeding units of speech back to the subject after a one-second delay. It was hypothesized that the procedure would increase the frequency of verbal behavior by adding to the auditory impressions the child gets from hearing samples of his own speech. The vocal responses of children under free-field feedback conditions were compared to responses under control conditions. Children of 'high' and 'low' verbal levels (PLS) were included and the interaction between the verbal levels and the experimental treatments was investigated.

The studies reported in Sections One, Two, and Three do not describe the full range of studies that have been completed at the Parsons Research setting. The studies presented do illustrate, however, the frame of reference used in the research setting and the directions in which the Project is moving. Subsequent reports will be necessary to complete the picture.

References

1. KANTOR, J. R., *An Objective Psychology of Grammar*. Evanston, Ill.: Principia Press, 1952.
2. MATTHEWS, J., Speech problems of the mentally retarded. Chapter 17 in Travis, L. E., *Handbook of Speech Pathology*. New York: Appleton-Century-Crofts, Inc., 1957, 531-551.
3. MILLER, N. E., and DOLLARD, J., *Social Learning and Imitation*. New Haven: Yale University Press, 1953.
4. MOWRER, O. H., The psychologist looks at language. *The American Psychologist*, 9, 1954, 660-694.
5. ROSENBERG, S. Methods for the rational assembly of aircrews. *J. aviat. Med.*, 28, 1957, 185-189.
6. ROSENBERG, S. Interpersonal processes in the perpetuation and reduction of language retardation: some speculations and some data. Paper given at American Association on Mental Deficiency, Milwaukee, 1959.
7. ROSENBERG, S., ERLICK, D., and BERKOWITZ, L. Some effects of varying combinations of group members on group performance measures and leadership behaviors. *J. abnorm. soc. Psychol.*, 51, 1955, 195-203.
8. SCHIEFELBUSCH, R. L. Establishing an interdisciplinary research team in a training school. *Am. J. ment. Defic.*, 65, 1960, 214-220.
9. SCHIEFELBUSCH, R. L., and BAIR, H. V. Joint setting for research: Training school and university. *Mental Hospitals*, Nov., 1960, 34-36.
10. SKINNER, B. F. *Science and Human Behavior*. New York: Macmillan Co., 1953.
11. SKINNER, B. F. *Verbal Behavior*. New York: Appleton-Century-Crofts, Inc., 1957.
12. SPRADLIN, J. E. Language and communication of mental defectives. Chapter 17 in Ellis, N. R,. *Handbook of Mental Deficiency*. New York: McGraw-Hill (in press).

Section One

Assessment of Speech and Language of Retarded Children: The Parsons Language Sample

JOSEPH E. SPRADLIN

Introduction

When one engages in speech and language research, he is immediately faced with measurement problems. The need for reliable and relevant measures is most apparent (a) when a sample is selected for an experimental study, (b) when the effects of experimental treatments are evaluated, and (c) when the interaction effects of treatment on type of subject are evaluated.

The evaluation measures most commonly used in the investigation of speech and language of children have been *age of the first word*, *vocabulary level*, *grammatical construction*, and *accuracy of articulation* (*1*). However, a review of these methods for evaluating speech and language in children reveals a lack of precision in estimating reliability and in determining equivalence of results obtained by different procedures. Moreover, none of these measures or evaluation procedures are systematically derived from, nor related to, a general behavioral system.

Such a behavioral system provides for more adequate evaluation of speech and language by focusing on important areas which would not be sampled in a less systematic approach. Likewise, the use of a system provides for an economy in systematic evaluation by indicating the situations in which the test constructor is sampling behaviors which are theoretically equivalent.

The behavioral system used in developing the Parsons Language Sample was drawn primarily from Skinner (*7*, *8*). Other behaviorists (*2*, *3*, *4*, *5*) have also presented systems for analyzing language behavior. The writer selected the Skinnerian model primarily because of its emphasis on the environmental conditions under which language behavior occurs.

The following discussion will be divided into (a) the presentation of a rationale for language assessment which is

Joseph E. Spradlin (Ph.D., George Peabody College, 1959) is Field Director of the Parsons Project and Research Associate in the Bureau of Child Research, University of Kansas. He wishes to thank Dr. John deJung, who supervised the statistical analysis of the data and served as a constructive and supportive critic throughout the preparation of the manuscript. He also expresses appreciation to Gregg Benjamin, William Finley, Teresa Hickey, and Dorothy Smith, who served as examiners.

related to a behavior system, (b) a description of the speech and language test derived from the rationale, and (c) empirical data relevant to the evaluation of the speech and language test and the conceptual system from which it was derived.

A Rationale for Language Assessment

The method used in obtaining a sample of language behavior in the PLS involved classification of language behavior and the preparation of items to sample language according to these several classifications. The guiding considerations used in the classification of language were (a) whether the language was vocal or non-vocal and (b) the conditions evoking or controlling its occurrence.

Classifications According to Whether Language is Vocal or Non-Vocal. Within the present text language may be either vocal or non-vocal. Vocal language is the product of the speech mechanism and is almost synonymous with speech. Non-vocal language always involves observable bodily movement, including facial movements.

Classification According to Controlling Conditions. Both vocal and non-vocal language are subject to the control of situations external to the speaker. The system for classifying language behavior according to the situations in which it occurs is drawn from Skinner (*9*). Three terms are used by Skinner in describing behavior. The terms are discriminative stimulus, response, and reinforcement. The *discriminative stimulus* is roughly equivalent to the stimulus context in which a given response will be reinforced. The *response* is an act or bit of behavior of the person. *Reinforcement* is the stimulus change or consequence which the response brings about. Reinforcement serves to change the frequency of occurrence of responses or behaviors. For instance, a person in a restaurant might say to the waiter, 'Please bring me a menu.' The waiter will usually oblige by bringing the menu. In this case the restaurant and waiter are the discriminative stimuli for the person's response, 'Please bring me a menu.' The reinforcement is the waiter's act of bringing the menu. We may consider the response, 'Please bring me a menu' as voluntary and not under stimulus control. However, note that we do not make this response to the postman since we have learned that the response does not lead to reinforcement under those conditions.

Skinner's major classes of language behavior based on this three-term contingency include mand, tact, and verbal behavior under the control of verbal stimuli.

Mand behavior includes such behavior as demanding, commanding, requesting, and asking. Usually the reinforcement for mand behavior is rather specific and it is often related to specific drive operations. If one is to increase the probability of a child saying, 'I'm thirsty,' which may be functionally equivalent to saying, 'Get me something to drink,' there are several things he might do. He might take the child for a long ride in a car, making certain that no water or other liquid was available to the child. This is a simple deprivation operation. Experimental observations indicate that an organism is more likely to emit a response which has previously been followed by

the delivery of water if it has not had water for several hours. The probability that the child would emit a mand response such as, 'I'm thirsty,' would also be greater if the child had previously been given salted peanuts. Giving the child salted peanuts and keeping him away from water would be classified as drive operations. Not all mands, however, are dependent on operations affecting physiological conditions. Another way of increasing the probability of a mand response is to increase the strength of other responses which are dependent on specific implements or conditions (which are not available) for their completion. For example, a child will probably request a pencil if asked to draw a picture but is given no writing or drawing implement. This request is related to an object which is necessary for the completion of another response. Mands may be either vocal or non-vocal (gestural). For example, the 18-month-old child pointing to the cookie jar while looking toward his mother is executing a gestural mand.

Tact behavior, unlike mand behavior, is primarily under the control of discriminative stimuli rather than under the control of specific drive operations. Certain responses will be followed by reinforcement only if they occur in the presence of specific discriminative stimuli. For example, the very young child's response 'cat' is most likely to be reinforced (praised) if the small four-legged animal which members of the community label 'cat' is present. Thus, the response 'cat' comes under the control of the discriminative stimuli presented by the cat. This naming response is what Skinner calls a 'tact.' Non-vocal tacts also occur. The iconic use of gestures to designate a given activity or object is often noted among children with speech and hearing handicaps. For example, a circular movement of the forefinger may occur in the presence of a wheel, or a back-and-forth motion may occur when a swing is involved as the discriminative stimulus for the response.

Verbal behavior under the control of verbal stimuli. One of the common types of such behavior is the *echoic response.* The echoic response is the repetition of a response that has been made by another person. That is, there is a point-to-point relationship between the discriminative stimulus presented by one person and the verbal response of a second speaker which follows it. A true echoic response is completely under the control of the verbal events that have immediately preceded it.

Echoic responses of children are reinforced rather consistently by parents and teachers. The infant who repeats his mother's 'bye, bye,' or 'ma, ma,' is reinforced by having his mother smile at him, talk to him, and cuddle him. When the verbal behavior of either children or adults is carefully observed, one is struck by the frequency of the echoic response.

Echoic gestures also play a role in the social situation. For example, most mothers spend considerable time in waving 'bye, bye,' at their child in order to teach him to wave 'bye, bye,' in imitation. Later the mother's waving behavior serves as a discriminative stimulus for the child's waving. It is apparent that echoic responses may be either vocal or non-vocal.

A second type of verbal response to verbal stimuli is *intraverbal behavior.* Intraverbal responses are responses which

are primarily under the control of verbal stimuli but have no point-to-point correspondence to them. For instance, one person's, 'Good morning, how are you?' may be a discriminative stimulus for, 'Fine, thank you.' Or, the person's own responses '2 + 2 =' is the discriminative stimulus for 'four.' In other words, stimuli produced by other verbal responses are the occasion for the intraverbal response. A given verbal response may be under the control of more than the single preceding word. In the statement, 'In the winter time the ground is covered with white _____,' the most probable response is 'snow.' If the discriminative stimulus were simply 'white,' then feathers, cotton, paper, or numerous other responses might follow equally as well. The words 'winter,' 'ground,' and 'covered' increase the probability of 'snow' being emitted.

Examples of intraverbal gestural responses are quite common. If a speechless child is asked a question such as, 'What do you do with a key?' his most probable response will be a twisting movement of his thumb and forefinger. This situation involves an intraverbal gestural response since the child is making a non-vocal response to a verbal stimulus. There is no point-to-point relation between the stimulus and the response.

Comprehension is a third type of response to verbal stimuli. Comprehension is a construct which is also based on differential responses, that is, on behavior emitted by the subject. We say a person comprehends what is being said to him when the person makes a differential response which bears some relation to a specific stimulus situation. The mode of stimulation may vary and we may thus make reference to auditory, visual, or tactile 'comprehension' depending on the stimulus conditions. The stimulus situation may be very complex and the response related to it very simple. Nevertheless, when we speak of comprehension, we are speaking of a relationship between some situation and some response. Comprehension is thus handled at the same level as other verbal behavior. Such a response may occur immediately following the stimulus or it may be delayed a considerable period of time, as is the case when a person follows the road directions given by a service station attendant.

Description of the PLS

The development of the PLS represents a systematic attempt to develop subtests and test items which sample language behavior according to the Skinnerian system. Seven subtests and 123 test items are used. The subtests of the PLS are *tact*, *echoic*, *intraverbal*, *echoic gesture*, *comprehension*, *intraverbal-gesture*, and *mand*. The tact, echoic, and intraverbal subtests all sample the child's vocal or speech behavior. The echoic gesture, comprehension, and intraverbal-gesture subtests were specifically designed to sample non-vocal communication. The mand subtest was designed to sample both vocal and non-vocal responses.

Description of Subtests and Administration Procedures. Testing is done in a small quiet room equipped with a desk or table and two chairs. The child is seated across the table from the examiner. The box of testing equipment is kept out of sight (under the table or in a desk drawer) of the child. After a piece of equipment has been used, it is

placed back in the box or drawer before the next item is administered unless the next test item utilizes the same equipment. Vocal responses to vocal subtest items are recorded verbatim. Responses to each subtest are scored according to whether the response is appropriate (√), inappropriate (×), unintelligible (※), or no response (—).[1] However, the empirical results presented in this chapter utilize only scores for appropriate responses.

Tact Subtest. The tact subtest consists of 28 objects or pictures which are named by the child. Seven items involve real objects, seven involve miniature objects, seven involve colored pictures, and seven involve non-colored pictures. The examiner shows the child each object or picture individually. Upon presenting each item the examiner says, 'What is it?' or 'What do you call it?' If the child does not respond, the examiner repeats the question. The examiner accepts and rates any response given by the child. After the child has responded, the examiner says, 'Good,' or indicates approval to the child and hands the child the toy for inspection. As soon as it is feasible, the examiner retrieves the object and proceeds to the next item. If there are no appropriate responses to any of the first five items, the examiner discontinues the tact subtest and proceeds to Mand Item 1. If any appropriate responses are given to the first five items, the examiner administers all 28 tact items. The responses which are classified as appropriate for each of the 28 tact items are shown in the test booklet in Appendix A.

Echoic Subtest. The echoic subtest consists of 22 items. The first ten items consist of words and sentences of varying degrees of complexity which the child is requested to repeat. Each word or sentence is introduced by the command, 'Say.' The 12 following items consist of a series of digits which the child is requested to repeat. The items range in difficulty from one to six digits —that is, two items require the child to repeat one digit, two items require him to repeat two digits, and so on through six digits. Prior to administering each item the examiner makes certain he has the child's attention. This precaution is necessary since the examiner presents the words, sentences, and numbers only once. The child's response is classified as appropriate if the words, sentences, or numbers are repeated as the stimulus is presented. Provided the child's speech is intelligible, errors in articulation are allowed. If any item is rated as appropriate, the following two items are administered in each section of the subtest. After three consecutive items are missed in the first section of the echoic subtest, that section of the subtest is discontinued and the examiner presents the items of the second section. After three consecutive items are missed in the second section of the echoic test, the examiner presents the second mand item.

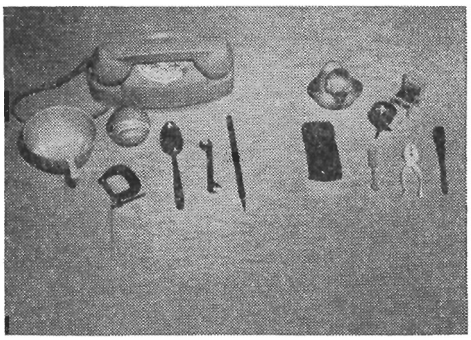
[1] An *appropriate response* to an item is a response that would be given by a large per cent of normal children or adults if they were presented the test item. A vocal response must be intelligible for it to be rated as appropriate. More specific criteria for rating responses as appropriate are given in the test booklet (Appendix A). An *inappropriate response* to an item is an intelligible response which would be given by only a very small per cent of normal children or adults if they were presented the item. The *unintelligible response* category applies only to vocal responses. These are nonconventional or inarticulate noises. The *no response* rating is given when the child makes no differential vocal or non-vocal response to the test item.

Echoic Gesture Subtest. The 13 items of the echoic gesture subtest sample the child's ability to mimic a motor act performed by the examiner. The items range from the examiner pointing to the light and saying to the child, 'Do this,' to having the child follow the examiner's motions as he taps two Kohs blocks in a complicated left-right sequence. This test measures the child's imitation of the non-verbal performance of another person. The appropriate response is demonstrated by the examiner three times on each item. If the child is successful in imitating the gesture on any of the three trials, he is given full credit. The examiner discontinues the subtest if the child fails two consecutive items from item nine through 13.

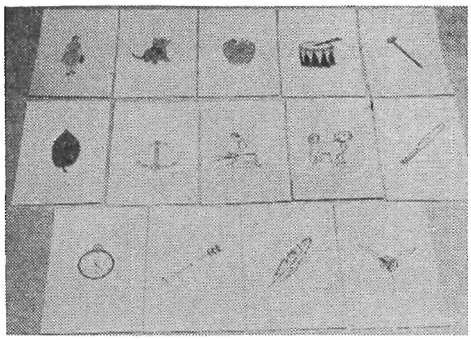
Comprehension Subtest. The comprehension subtest consists of a series of 18 items in which the examiner directs the child to complete certain motor tasks. Five of the items are given with vocal directions only, five are given with gestural directions only, and five are given with a combination of both vocal directions and gestures. With the exception of the first item, in which the examiner says the child's name and determines if the child responds, the child's attention is obtained prior to the administration of the item. If the child responds (appropriately or inappropriately) the first time the item is administered, the directions are not repeated. If the child does not respond the first time the directions are administered, they are repeated once. The directions are repeated exactly as given originally. The response to the first item in which the examiner says the child's name is considered appropriate if the child looks at the examiner. For the remaining items the response is classed as appropriate only if the child executes the command in the proper sequence. The comprehension subtest is discontinued after five consecutive failures (no response or inappropriate response).

Intraverbal Subtest. The intraverbal subtest is composed of 29 items which sample the child's vocal responses to vocal stimuli. For the first seven items the examiner asks the child simple questions such as, 'What do we do when we are hungry?' The next 16 items present the child with an incomplete sentence such as, 'The flag is red, white, and _____.' The sentences are presented orally by the examiner. The final six items are questions concerning similarity such as, 'In what way are a cat and a dog alike?' Prior to reading each question to the child the examiner obtains the child's attention. If the child does not respond, the question is read again. If the child responds, his response is recorded and the question is not read again. If none of the responses to the first five items of the intraverbal subtest are scored as appropriate, the subtest is discontinued. If any of the responses to the first five items are appropriate, all 29 items are administered. The appropriate responses to the items of the intraverbal subtest are given in the test booklet in Appendix A.

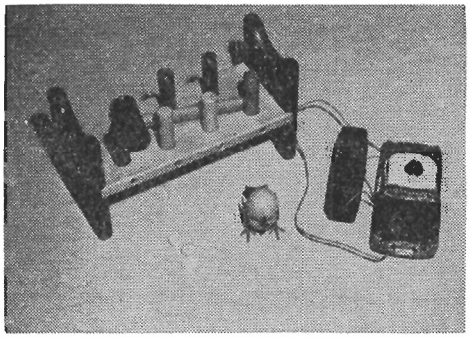
Intraverbal Gesture Subtest. The intraverbal gesture subtest consists of 24 questions which can be answered with either a verbal or a gestural response. Some questions such as item 2, which asks, 'Where is your ear?' very frequently elicit gestures whereas others such as item 24, which asks, 'What do you do with a handkerchief?' are less likely to elicit gestures. While vocal responses are also elicited, only gestures



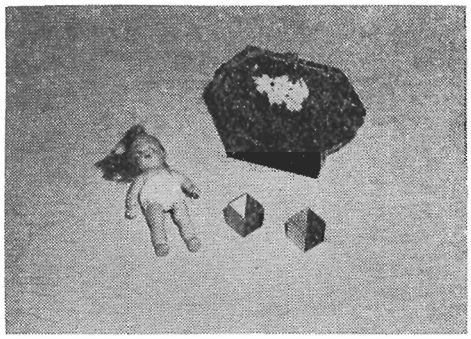
Objects for Tact and Comprehension subtests



Picture materials for Tact subtest



Objects for Mand subtest



Objects for Comprehension and Echoic Gesture subtests

FIGURE 1.1. Materials used in administering the tests constituting the Parsons Language Sample.

are scored.[2] Prior to asking the child the question the examiner is to make sure he has the child's attention by tapping the table or saying the child's name. If the child makes no appropriate vocal or gestural response to the first five items, the subtest is discontinued. Otherwise all 24 items are administered.

Mand Subtest. The mand subtest consists of five items interspersed among the other subtests. Both vocal and gestural responses are rated on four of the five items. On the fifth item, only the vocal response is rated. As an example, for item 1 the examiner holds a brightly colored toy wind-up duck in the child's view. If the child does not request the object, the examiner winds the duck up and places it on the table out of the child's reach for about five seconds. Finally, if the child still does not request the object, the examiner picks up the object, holds it in the child's view, lets it run for five seconds, and then places it back in the box or drawer. An appropriate mand response is scored if the child requests the

[2]Initially the aim of the intraverbal gesture was to measure both vocal and gestural behavior. However, during the initial testing one examiner focused on vocal responses and failed to record gestural responses. At this point examiners were instructed to score only the gestural responses.

object at any time by a gesture, speech, or both. The examiner scores for both speech and gestures on this item. The crucial aspect of scoring a mand item involves the type of response a child makes when conditions are set up whereby the child 'attempts' to obtain an object from the examiner. If the child asks for the object or gestures for it, the examiner records either an appropriate vocal or gestural mand. If he merely sits or reaches for the object, a no response or inappropriate response is recorded.

The materials required for administering the test are shown in Figure 1.1.

Empirical Evaluation of the PLS

The PLS was initially administered to 275[3] ambulatory mentally retarded children between the ages of seven years, 11 months and 15 years, eight months at the Parsons State Hospital and Training Center. Four persons without previous experience in administration of psychometric tests were used as examiners. Two of the examiners were male students at the Parsons Junior College. The other two were housewives who had had college training. None of the examiners had had previous experience with mentally retarded children.

All four examiners were told that the Project Staff was developing an instrument to evaluate children's vocal and non-vocal language and that it was necessary to administer the test in order to evaluate it. After a brief discussion of the purpose of the PLS the following training procedure was initiated. The test was broken into two sections consisting of the first three and last three subtests. Mand items which came between the first three subtests were learned as a part of the first portion of the test. Mand items which came during the second section were learned when that section was administered. Each examiner first observed the writer through a one-way mirror as he administered the first three subtests. The writer then discussed the test administration with the examiner and answered any question which occurred concerning test administration. The examiner was then given a test kit and a test blank and asked to go home and study the first three subtests, which he would be required to administer on the following day. The following day the examiner administered the first three subtests to two different children while the writer observed. After each examination the writer discussed the test administration with specific attention to the examiner's rigor in (a) following test instructions for administration, (b) recording, and (c) excluding remarks unrelated to test administration.

The procedure used in training the examiner to administer the first three subtests was then repeated for the second three subtests. After the examiner had administered each part twice, the writer administered the complete test while the examiner observed. The test administration was discussed and the examiner then administered the complete test. If the test administration was adequate, the examiner was assumed to be prepared to administer the test. If it was not adequate, the errors in administration were discussed and the examiner was again observed as he administered the complete test. The judgment of whether the examiner administered the

[3]Due to examiner error, scores for 29 subjects on the intraverbal gesture subtest were not usable. See footnote 2 above.

TABLE 1.1. Maximum differences in cumulative proportions and Kolmogorov-Smirnov tests for significance between pairs of examiners for each subtest.

Subtest	*Examiners*	*N*	*Examiners* 1	2	3
Tact	1	76			
	2	58	.071		
	3	83	.166	.130	
	4	58	.099	.088	.182
Echoic	1	76			
	2	58	.083		
	3	83	.127	.067	
	4	58	.088	.052	.045
Intraverbal	1	76			
	2	58	.106		
	3	83	.134	.110	
	4	58	.083	.086	.139
Echoic Gesture	1	76			
	2	58	.141		
	3	83	.108	.079	
	4	58	.138	.102	.069
Comprehension	1	76			
	2	58	.160		
	3	83	.184	.119	
	4	58	.035	.034	.084
Intraverbal Gesture	1	47			
	2	58	.175		
	3	83	.115	.087	
	4	58	.225	.121	.139
Mand Vocal	1	76			
	2	58	.211		
	3	83	.329*	.118	
	4	58	.455*	.310*	.285*
	1	76			
	2	58	.454*		
	3	83	.329*	.285*	
	4	58	.210	.310*	.552*

*These differences in proportions are significant at the .05 level since they are greater than the .272 value necessary for significance when $N_1 = 47$ and $N_2 = 58$. All other proportion differences are less than the .218 value necessary for significance when $N_1 = 76$ and $N_2 = 83$.

test adequately was based on correct reading of items, limiting help to that allowed in the test directions, correct placement of objects on the desk, proper use of gestures (used when called for, not used when not called for), and timing.

Examiner Equivalence. An immediate concern was that of examiner equivalence. Do different examiners obtain equivalent results when using the PLS? If the PLS is to have utility as a research instrument, the test scores must be primarily a function of the examinees' behavior rather than a function of the mode of item presentation or scoring of the examiner. To explore this question, frequency distributions were obtained for each of the four examiners for each subtest. The subtest score distributions

TABLE 1.2. Means, standard deviations, and F values for differences on the measures of the PLS.

	Examiner 1			*Examiner 2*			*Examiner 3*			*Examiner 4*				*F Maximum*
	$\bar{X}$	*SD*	*N*	$\bar{X}$	*SD*	*N*	$\bar{X}$	*SD*	*N*	$\bar{X}$	*SD*	*N*	*F Value*	*Value*
Tact	13.50	9.39	32	16.17	5.63	42	16.35	8.09	69	15.97	8.09	45	1.07	2.78*
Echoic	9.34	6.35	32	11.14	4.98	42	10.03	5.65	69	11.22	5.20	45	1.07	1.63
Intraverbal	9.09	9.43	32	9.57	8.47	42	11.22	9.22	69	10.87	9.06	45	.57	1.24
Mand Vocal	.53	.62	32	.80	1.02	42	1.35	.74	69	1.00	1.48	45	5.44*	5.76*
Vocal	31.94	24.24	32	36.88	17.59	42	37.59	21.45	69	38.07	20.88	45	.65	1.90
Comprehension	8.27	4.94	37	10.30	5.00	53	10.09	4.81	80	10.11	4.92	53	1.54	1.08
Echoic Gesture	5.84	3.65	37	6.72	2.96	53	6.06	3.29	80	6.36	3.45	53	.64	1.52
Intraverbal Gesture	5.08	4.88	37	7.04	5.68	53	5.33	5.13	80	6.42	5.74	53	1.53	1.38
Mand Non-Vocal	.19	.62	37	.68	.89	53	.79	.88	80	1.28	1.09	53	10.93*	3.13*
Non-Vocal	19.19	11.96	37	24.06	11.00	53	21.48	10.96	80	22.89	11.64	53	1.50	1.19

*Significant at the .05 level.

were based on the scores from every child which the examiner tested. Kolmogorov-Smirnov two sample tests (*9*) which are sensitive to differences in central tendency, skewness, and variability were used to test for difference in score distributions among the four examiners. The distribution of each examiner's scores was compared with that of every other examiner's scores for each subtest. The results of the Kolmogorov-Smirnov tests shown in Table 1.1 indicate that only the score distributions for the two mand subtest measures yield significant differences between examiners.

Examiner differences in mean subtest scores for each of the subtests and for the vocal and the non-vocal composite subtest scores were examined using an analysis of variance design. The vocal composite score is a simple sum of the three vocal subtests, the tact, echoic, and intraverbal subtests. The non-vocal composite score is the sum of the three non-vocal subtests—echoic gesture, comprehension, and intraverbal gesture. Subjects who scored zero on all of the individual subtests of either the vocal or the non-vocal composite were designated as untestable for that particular test section and for all subtests within that section. Eighty-seven of the 275 subjects scored zero on the vocal subtests; 51 subjects scored zero on the non-vocal subtests. The distribution of chronological ages for these 'zero score' subjects was roughly similar to the age distribution for those retained. The means, standard deviations, and F values for each subtest for each examiner are presented in Table 1.2.

Differences in subtest score variance between examiners for each of the subtests and for the two composite subtests were examined by means of the Harley Test (*9*). The maximum F ratios are reported in the last column in Table 1.2. Only the ratios for the two mand subtests and for the tact subtest were found to exceed the values for the .05 level of confidence, suggesting the untenability of the hypothesis of homogeneity of

TABLE 1.3. Spearman rank order correlation coefficients computed between examiners for each of six PLS subtests.

Tact (N = 28 items)					*Echoic (N = 22 items)*					*Intraverbal (N = 29 items)*				
Examiners					*Examiners*					*Examiners*				
	1	*2*	*3*			*1*	*2*	*3*			*1*	*2*	*3*	
2	.94				2	.96				2	.94			
3	.96	.98		Mdn = .95	3	.98	.96		Mdn = .97	3	.92	.96		Mdn = .93
4	.94	.94	.97		4	.98	.98	.97		4	.92	.92	.96	
Echoic Gesture (N = 13)					*Comprehension (N = 18)*					*Intraverbal Gesture (N =24)*				
Examiners					*Examiners*					*Examiners*				
	1	*2*	*3*			*1*	*2*	*3*			*1*	*2*	*3*	
2	.95				2	.96				2	.95			
3	.92	.92		Mdn = .94	3	.95	.98		Mdn = .96	3	.78	.76		Mdn = .81
4	.93	.93	.97		4	.94	.99	.97		4	.84	.90	.67	

variance between examiners for these three subtests. Examiner differences in subtest score distribution are within chance variability for the remaining subtests.

F ratios for differences in subtest score means between examiners are reported in the last column of Table 1.2. The hypothesis of no differences between examiners in subtest means is rejected for the two mand tests at the .05 level. Examiner differences for all other subtest score means are well within chance variability.

A more direct analysis of whether different examiners obtain similar results with the PLS was made in terms of the percentages of subjects making an appropriate response for different examiners on each item in each subtest.[4] For each examiner within each subtest the items were ranked in terms of these percentages of appropriate response and Spearman rank order correlation coefficients computed for each pair of examiners for each of the six subtests. These coefficients are presented in Table 1.3.

All of the rank order coefficients between examiners are above .90 for the tact, echoic, intraverbal, echoic gesture, and comprehension subtests. The median correlations for these tests are .95, .97, .93, .94, and .96, respectively. The coefficients for the intraverbal gesture subtest range from .67 to .94 with a median correlation of .81. The somewhat lower correlations for the intraverbal gesture subtest indicate that examiners are more varied in their scores for responses on this subtest than they are on the other five; evidently they are eliciting gestures differently on the various items or are scoring them differently when they occur. Closer examination of the inter-examiner coefficients indicates that most of this lower agreement is associated with examiner number three. Generally the other three examiners highly agree.

The generally high rank order coefficients presented in Table 1.3 say nothing about constant differences between examiners in percentage of items passed. However, since there were no significant differences found in terms of the Kolmogorov-Smirnov tests of equivalent percentages of subjects achieving different subtest scores nor in the F ratios for subtest score means, the possibility of

[4]The vocal and non-vocal mand subtests contain only four and five items, respectively, and were therefore not included in this analysis.

TABLE 1.4. Odd-even reliability coefficients for the vocal subtests (N = 188) and for the non-vocal subtests (N = 224) of the PLS.

	VOCAL				*NON-VOCAL*			
No. of Items	*Tact*	*Echoic*	*Intra-verbal*	*Vocal Mand*	*Echoic Gesture*	*Compre-hension*	*Intra-verbal Gesture*	*Non-Vocal Mand*
r 1/2 I/II	.94	.91	.96	.25	.86	.87	.84	.17
r 1/2 I/II (corrected)*	.97	.95	.98	.40	.92	.93	.91	.29

*Using Spearman-Brown formula for reliability of the whole test.

constant examiner differences may be discounted.

In summary, given the examiner training procedures outlined above, the data from all these examinations of inter-examiner difference suggest lack of examiner bias in terms of the administration and scoring of subject responses for six of the PLS subtests. On the other hand, inter-examiner differences were noted for the two mand subtests.

Reliability. An initial examination of the reliability of the several PLS subtests was made in terms of split half coefficients computed from data obtained from the PLS administration to the 275 subjects described above. As in previous analyses only scores for subjects receiving other than zero scores on one or more of the vocal subtests (N = 188) and other than zero scores on one or more of the non-vocal subtests (N = 224) were included in these computations. See Table 1.4.

The low split half coefficients were obtained for the two mand subtests, correlations of .25 for the vocal scoring and .17 for the non-vocal scoring.[5] The coefficients are all above .90 for the three subtests of the vocal section of the PLS. The correlations for the three subtests of the non-vocal composite section are all above .84. When the Spearman-Brown correction formula for double test length was applied, the reliability coefficients for the three vocal subtests were all above .95. The Spearman-Brown coefficients for the three non-vocal subtests were all above .91.

A further examination of subtest reliability was made in terms of test-retest data for a 40-subject sample of the original 275 subjects. These 40 children were retested by one of the original four examiners. Twenty children were drawn randomly from the children she had seen previously and 20 children were drawn randomly from the children who had been seen previously by the other three examiners. The period of time between the first tests and the retests ranged from two to five months. Only scores for subjects classified as testable (other than only zero scores on one or more PLS subtests of either the vocal or the non-vocal sections on one of the two test administrations[6]) were retained for

[5] In view of these very low split half reliability coefficients and the poor examiner equivalence mentioned previously, the two mand subtests were deleted from this and all subsequent subtest analyses.

[6] If a subject was testable on one test but untestable on the second, he was given the lowest possible score on the various measures of the test on which he was untestable and his data were included in the analyses.

TABLE 1.5. Means, standard deviations, *t* values, and correlation coefficients for subtest and composite PLS scores obtained by the same and different examiners retesting 40 subjects two to five months after the initial testing.

Measure	*Conditions*		*N*	*Mean*	*SD*	*D*	*t*	*r*
			VOCAL SUBTESTS					
Tact Subtest	Same Examiner	Test	15	49.27	9.96	1.40	1.19	.90
		Retest		50.67	9.24			
	Different Examiners	Test	15	49.13	11.22	2.00	1.76	.93
		Retest		51.13	9.59			
Echoic Subtest	Same Examiner	Test	15	47.93	10.15	3.27	2.27*	.86
		Retest		51.20	9.98			
	Different Examiners	Test	15	50.60	10.53	1.40	1.25	.92
		Retest		52.00	10.19			
Interverbal Subtest	Same Examiner	Test	15	48.53	8.80	.40	1.04	.99
		Retest		48.93	8.97			
	Different Examiners	Test	15	51.13	9.72	.67	1.61	.99
		Retest		51.80	9.99			
Total Vocal Section	Same Examiner	Test	15	145.73	27.55	5.07	2.08	.94
		Retest		150.80	26.84			
	Different Examiners	Test	15	150.86	29.23	4.06	2.13	.97
		Retest		154.93	27.01			
			NON-VOCAL SUBTESTS					
Echoic Gesture Subtest	Same Examiner	Test	19	50.10	9.98	1.16	.61	.64
		Retest		51.26	8.81			
	Different Examiners	Test	20	44.65	11.68	2.40	4.28*	.88
		Retest		47.05	11.20			
Comprehension Subtest	Same Examiner	Test	19	51.63	12.17	−.68	.43	.84
		Retest		50.95	8.95			
	Different Examiners	Test	20	49.25	11.72	.25	.24	.92
		Retest		49.50	11.09			
Intraverbal Gesture Subtest	Same Examiner	Test	19	52.26	7.32	.16	.11	.73
		Retest		52.42	9.32			
	Different Examiners	Test	20	47.70	8.50	−.20	.16	.77
		Retest		47.50	6.70			
Total Non-Vocal Section	Same Examiner	Test	19	154.00	24.67	.63	.19	.73
		Retest		154.63	21.82			
	Different Examiners	Test	20	141.60	26.66	2.45	1.43	.96
		Retest		144.05	25.63			
			TOTAL PLS					
	Same Examiner		19	305.07	47.61	8.06	1.28	.89
				313.13	33.31			
	Different Examiners		20	303.27	47.60	7.40	2.41*	.98
				310.67	41.89			

*Significant at the .05 level.

this temporal stability analysis. The pairs of test-retest scores based on repeated administrations by the same examiner (N = 15 for the vocal subtests and 19 for the non-vocal subtests) and those pairs of scores involving an administration by a different examiner (N = 15 for the vocal subtests and 20 for the non-vocal subtests) were analyzed separately in terms of matched *t* ratios and product moment correlation coefficients. A summary of this analysis is presented in Table 1.5.[7]

As may be seen in Table 1.5, the retest scores are somewhat higher than those for the initial test, both for readministration by the same examiner and for readministration by a different examiner. However, when both the test and retest were administered by the same examiner, only the gain for the echoic subtest was significant at the .05 level. Only the gains for the echoic gesture and PLS total scores were significant when different examiners administered the test and retest. A reasonable explanation here might be that language improved or that the examiner changed during the intervening period of two to five months. The fact remains, however, that only three of 20 possible gain increments were significantly different from zero. It is apparent that additional data are needed before further interpretations can be made.

The high test-retest product moment correlations for all the vocal subtest scores and, to a lesser degree, for the non-vocal subtest scores, present an encouraging picture of the stability of the examiners' relative rank ordering of the subjects' subtest scores over a several month period. Further support for the earlier data (pages 16-17 above) of examiner equivalence comes from the fact that coefficients based on different examiners very nearly approximate those for the same examiner.

Test data from 32 children were used in examining the temporal stability of the PLS. The children were those between nine and 12 years of age who scored between the 27th and 73rd percentiles of that age group on the total score of the PLS. These 32 subjects were among the 275 children initially tested on the PLS. The retesting was administered by the same examiner who gave the first test. The length of time between the first and second testing was approximately 29 weeks. Only subjects receiving other than zero scores on both the initial and subsequent administration of a given subtest were included for analysis of that subtest. Test-retest correlations were computed for each subtest using raw scores and for subsection and total scores using summated standardized scores. Means were compared by *t* ratios. Results are presented in Table 1.6.

Although four of the six subtests showed a mean gain over the 29-month

[7] A conversion table was used to equate the means and standard deviations of scores for each of the retained six PLS subtests. The conversions were made in order to abrogate differences in number of items, difficulty level, and differential variability. Subtest means and standard deviations were set at 50 and 10, respectively. The conversions were made utilizing all usable subject scores from the initial test sample of 275 subjects, 188 usable subjects for the vocal subtests and 224 usable subjects for the non-vocal subtests. The conversion table for all subtests is presented in Appendix B. The measures used to select and classify subjects in the studies reported in Sections Two and Three were derived from T-scores. The conversion tables for the T-scores are in Appendix C. Correlations between T-scores and standard scores for the PLS vocal, non-vocal, and total all were .96 or higher.

TABLE 1.6. Means, standard deviations, correlation coefficients, and *t* ratios for PLS test-retest data obtained by the same examiner retesting 32 children between nine and 12 years of age 29 months later.

	N†	Mean	SD	D	t	r		N	Mean	SD	D	t	r
Tact							*Echoic Gesture*						
Pre	23	12.43	4.34	.26	.45	.82	Pre	31	5.52	1.88	.42	1.27	.59
Post	23	12.17	4.70				Post	31	5.10	2.07			
Echoic							*Comprehension*						
Pre	22	8.18	3.64	.37	.74	.83	Pre	32	7.97	4.03	.22	.61	.89
Post	22	8.55	4.04				Post	32	8.19	4.69			
Intraverbal							*Intraverbal Gesture*						
Pre	9	5.89	5.51	.08	.17	.96	Pre	25	7.20	4.98	.96	1.33	.77
Post	9	6.89	5.43				Post	25	8.16	5.32			
Vocal							*Non-Vocal*						
Pre	23	131.22	17.19	2.87	2.09*	.93	Pre	32	145.00	23.92	.56	.32	.92
Post	23	134.09	17.07				Post	32	145.56	24.26			

Total

	N	Mean	SD	D	t	r
Pre	23	281.70	28.38	3.47	1.41	.93
Post	23	285.17	31.36			

*Significant at the .05 level.
†The N's vary considerably between subtests since only subjects who obtained scores other than zero on both administrations of a particular subtest were included.

interval, only the gain for the vocal composite score is significant at the .05 level. As with the previously discussed retest data, the general lack of significant increases in mean scores does not indicate change or growth in language behaviors.

The correlational data shown in Table 1.6 once again indicate that the measures of the PLS have generally high retest stability even on a restricted sample. Subtest test-retest correlations range from .82 to .96 for the vocal subtests and from .59 to .89 for the non-vocal subtests. The reliability coefficients for the composite vocal, composite non-vocal, and total PLS are .93, .92, and .93, respectively.

A further examination of temporal stability involved initial and follow-up testing of new male patients. Twenty-seven newly admitted boys between six and 15 years of age were tested from two to 62 days after their arrival at the Parsons State Hospital and Training Center. They were retested approximately 15 months later. The same examiner administered both tests. These boys were initially assigned to a single cottage. However, as time progressed, several of them were assigned to different cottages because they presented a different type of management problem from that of the other boys in the cottage.

All subtest raw scores were converted to standardized scores and test-retest correlations were computed for each subtest and subsection score and for the total PLS score. One subject who failed

TABLE 1.7. Means, standard deviations, *t* ratios, and correlation coefficients for test-retest data obtained by the same examiner on male new admissions after 15 months.

	N	*Mean*	*SD*	*t*	*r*		*N*	*Mean*	*SD*	*t*	*r*
Tact						*Echoic Gesture*					
Pre	25	49.81	9.60	.59	.68	Pre	27	49.22	7.07	1.63	.15
Post	26	48.88	10.11			Post	27	52.33	7.80		
Echoic						*Comprehension*					
Pre	26	49.69	8.58	.47	.62	Pre	27	54.63	8.66	.65	.59
Post	26	50.38	8.36			Post	27	53.70	7.14		
Intraverbal						*Intraverbal Gesture*					
Pre	26	49.23	8.86	.23	.74	Pre	27	52.79	10.57	.77	.50
Post	26	49.54	9.38			Post	27	54.37	10.44		
Vocal						*Non-Vocal*					
Pre	26	148.72	25.52	.05	.73	Pre	27	156.63	20.40	.96	.43
Post	26	148.90	26.68			Post	27	160.79	20.74		

Total

	N	*Mean*	*SD*	*t*	*r*
Pre	26	307.08	39.59	.49	.55
Post	26	310.81	40.81		

to score on any vocal subtest was dropped for that section of the analyses. *t* ratios were computed for all mean comparisons. These analyses are summarized in Table 1.7.

Only small non-significant differences were found between initial and follow-up subtest scores. Four of the six subtests showed a mean gain over the 15-month interval and two showed a slight loss. This pattern of gain and loss is consistent with that obtained in the previous follow-up study (Table 1.6). The mean gains for the subsection scores and for the total PLS score reflect the slight predominance of gains of subtests. None of these cumulative gains was significantly different from chance expectation at the .05 level.

The test-retest correlations shown in Table 1.7 are, in general, much lower than those shown in Tables 1.5 and 1.6 and range from a high of .74 to a low of .15. The lower correlations may be a function of such variables as an increased length of time between the two tests, the effect of individual differences in initial impact of the institution on the language behavior, or differential effects of prolonged institutionalization.

In summary, six subtests exhibit substantial score stability both in terms of within-administration stability (correlation of half scores) and in terms of subsequent administration by either identical or different examiners. As was found for the examiner equivalence data, the two mand subtests appeared distinctly less stable than did the remaining six subtests. The lower retest correlations involving newly admitted patients retested more than a year later suggest more hesitant predictive statements regarding such subjects.

Intercorrelation of the PLS Subtests and Subsections. It is generally desirable

TABLE 1.8. Correlation matrix for the subtests and subsections of the PLS for all testable subjects (N = 187).

	Mean	*SD*	*Var. No.*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
CA	153.27	24.60	1	.....								
Tact	50.26	10.02	2	.26	.....							
Echoic	50.22	10.11	3	.24	.77	.....						
Intraverbal	50.33	9.89	4	.32	.82	.81	.....					
Vocal	150.80	27.98	5	.30	.93	.93	.94	.....				
Echoic Gesture	52.36	8.16	6	.22	.64	.63	.65	.69	.....			
Comprehension	52.65	7.52	7	.23	.79	.69	.74	.80	.62	.....		
Intraverbal Gesture	51.37	8.83	8	.00	.12	.01	.02	.04	.16	.32	.....	
Non-Vocal	156.37	18.52	9	.19	.67	.56	.58	.64	.75	.85	.68	.....
Total	307.18	42.34	10	.28	.91	.85	.88	.94	.78	.90	.32	.86

for the subtests within any given test to have low to moderate interrelationships. Highly correlated subtests, in effect, duplicate their predictive utilities and serve principally as test lengtheners. The gain in validity from doubling or otherwise increasing the number of items in a test is of minor consequence if the first set of items yields reliable scores (*1*).

The correlations among the various subtests have implications concerning the utility of the rationale on which the PLS is based. The subtests were based on a classification of language behavior according to the controlling stimuli and the type of response occurring. A study of the relationship between the subtests will determine whether or not there is a large percentage of children who respond appropriately to one class of stimuli but not to other classes. Moreover, such a study should indicate whether there is a large percentage of children who respond appropriately at the motor level and not at a vocal level or vice versa. If there is a large percentage of such children, the correlations between the subtests should be quite low; if there is a small percentage of such children, correlations between subtests should be high. The relations between the various subtests and subsections of the PLS were examined both as a practical matter of prediction and in order to evaluate the utility of the rationale on which the subtests were originally constructed.

The examination of the PLS subtest and subsection score intercorrelations involved the 187 subjects (93 boys and 94 girls) of the initial PLS sample who were identified as testable on both the vocal and non-vocal subsections of the PLS.[8] Product moment correlations between the subtest and subsection scores[9] of the PLS were computed for these 187 subjects in order to examine the interrelationships between these measures. Initially all correlations were computed separately for boys and girls. Examinations of these two sets of correlations re-

[8]Those who obtained other than zero scores within both vocal *and* non-vocal section. Actually only one subject who was testable on the vocal subsection failed to make any appropriate responses on all gestural subtests.

[9]The subtest scores were those previously standardized on all testable subjects for any given subtest (see footnote 7 above). The subsection and total scores were computed as sums of these standardized scores.

vealed nearly identical intersubtest correlation coefficients for the two subgroups. The sex subgroups were then combined and a matrix of intercorrelations for the total 187 subjects computed. These coefficients are presented in Table 1.8.

The intercorrelations among five of the subtests (tact, echoic, intraverbal, echoic gesture, and comprehension) are moderately high, ranging from .62 between the echoic gesture and comprehension subtests to .82 between the tact and intraverbal subtests. The median for the ten correlations between pairs of these five subtests is .70. These correlation values mean that the five subtests have from 38 to 67% variance in common. The relatively high amounts of variance which these five subtests have in common indicate that when a vocal or non-vocal response is brought under the control of one type of discriminative stimuli (for example, visual) it can very likely be brought under the control of other discriminative stimuli (for example, auditory). Moreover, these findings also indicate that if a child exhibits appropriate vocal behavior, he will also exhibit appropriate non-vocal behavior.

The intersubtest correlations involving the intraverbal gesture subtest are all very low, ranging from .01 to .32. Scores on this subtest are independent of those from any of the other subtests. As previously reported in Table 1.4, moderately high odd-even half correlation of .84 for the intraverbal gesture subtest suggests a reasonable within subtest homogeneity. This suggests that there are similarities among items within the subtest since they elicit related subject responses. The high inter-rater agreement across random pools of subjects (the cumulative frequency comparisons of Table 1.1), across item difficulties (Table 1.3), and over time (Tables 1.5, 1.6, and 1.7) further demonstrates the objectivity and temporal stability of the scores of the intraverbal gestural subtest.

All subtests correlate substantially with the summated subsection of which they are a part and, except for the intraverbal gestural subtest, with the total PLS score. These part-whole correlations suggest that the vocal subtests could be combined with minor loss of prediction. The fact that the intraverbal gesture subtest is relatively objective and reliable and does not correlate with the other five subtests raises some interesting questions, such as, 'Are scores on the intraverbal gesture subtest related to the use of gestures outside of the test situation?' 'What is the function of gestures in communication among mentally retarded children?' and 'Under what conditions does the use of gestures develop?'

It is possible that all of the subtests, excepting the intraverbal gesture subtest, could be combined into a single measure without very much loss of predictive power. This, of course, is not the case with the intraverbal gesture subtest, which is independent of the other five.

The intercorrelation of .64 between the vocal and non-vocal subsections suggests that these two scores may have descriptive and predictive utility; however, it must be remembered that this correlation would, no doubt, be higher if the intraverbal gesture subtest were excluded from the non-vocal subsection.

Relationship between PLS Scores and Non-Language Test Behavior. The analyses of the PLS data presented thus far have been concerned with examiner

TABLE 1.9. Means and standard deviations for age, WB IQ's, and PLS scores for the older and younger groups.

Measures	*Older (N = 38)*		*Younger (N = 35)*	
	Mean	*SD*	*Mean*	*SD*
PLS Vocal	58.74	4.49	56.34	3.41
PLS Non-Vocal	55.39	3.22	56.54	3.27
PLS Total	57.05	3.02	56.66	4.86
WB-V	56.45	10.46	58.63	14.45
WB-P	57.63	15.02	55.74	11.65
WB-FS	54.50	11.53	58.74	10.46
Age	15.16	.81	11.52	1.48

equivalence, reliability (both in terms of single and repeated test administrations), and subtest intercorrelations. An equally important consideration is how the PLS scores relate to non-PLS behavior. To date, three types of studies have been completed which bear on this question. First, the PLS vocal, non-vocal, and total scores have been correlated with IQs of the Wechsler Scales. Second, the relationship of the PLS vocal and non-vocal scores with ratings of children's language behavior in non-test situations have been examined. Third, a series of studies has been conducted in which subjects of experiments were classified as 'high' and 'low' according to the vocal subtests of the PLS. These latter studies are reported in the next two sections of the Monograph.

Horowitz investigated the relationship between PLS subsection scores and verbal (WB-V), performance (WB-P), and full scale (WB-FS) IQs of the Wechsler Intelligence Scales (Wechsler-Bellevue, Wechsler Adult Intelligence Scale, and Wechsler Intelligence Scale for Children). One hundred and fifty-seven subjects in the institution had been tested by both the PLS and one of the Wechsler Scales. These subjects were divided into two age groups, 8-13 and 14-16, hereafter referred to as younger and older. Thirty-five children were drawn randomly from the younger group and 38 children were drawn randomly from the older group. The time interval between PLS testing and Wechsler Scale testing was no longer than two years for any subject.

Table 1.9 shows the means and standard deviations for age and test scores for each group. Pearson product moment correlations were computed within each age group for each of the following pairs of scores: WB-FS IQ and PLS total scores; WB-V IQ and the PLS vocal scores; and WB-P IQ and PLS non-vocal scores. Since conversion scores for the Wechsler Scales are designed to reduce variance associated with chronological age, a more direct comparison of the Wechsler Scale IQ scores and the PLS scores is made by partialling out age. The zero order coefficients and the coefficients with age partialled out are shown in Table 1.10. There was little difference between the zero order correlations and the partial correlations for any of the measures except for WB-V IQ and PLS vocal for the younger group. For these two measures the zero order correlation was .57 while the partial correlation was .68. As can be seen in Table 1.10, the correlations between the WB-V IQs and the PLS vocal scores

TABLE 1.10. Correlations between WB IQs and PLS measures for the older and younger groups.

	Older (N = 38)		*Younger (N = 35)*	
	r	*r partial*	*r*	*r partial*
WB-V–PLS Vocal	.74	.74	.57	.68
WB-P–PLS Non-Vocal	−.01	−.01	.15	.13
WB FS–PLS Total	.51	.51	.41	.41

are .68 and .74, respectively, for younger and older groups when age is partialled out. These two tests have from 46 to 55 common variance. The partial correlations between the WB-P IQs and the PLS non-vocal scores are near zero, being −.01 for the older group and .13 for the younger group. These two tests have negligible variance in common. The partial correlations between the PLS total scores and the WB-FS IQs are .51 and .41, respectively, for the older and younger groups indicating that these two tests have from 17 to 26% common variance.

In general, the moderately low correlations between the PLS scores and the Wechsler IQs indicate that these measures may be sampling different behavior within the population studies. This is especially true for the non-vocal sections of the two tests. Insofar as these measures of the PLS have generally proved reliable, this finding would indicate that the PLS may have predictive utility which is unassociated with the predictive utility of the Wechsler Scales.

In a second validation study the relationship between PLS vocal and non-vocal subsection scores and children's language behavior in the cottage situation was studied by having psychiatric aides in five Hospital cottages rank children of their respective cottages for speech and non-speech communication. The primary concerns of the study were the relationships between the PLS vocal score and the aide ranks for speech communication and the relationship between PLS non-vocal score and aide ranks for non-vocal communication.

The aides were first instructed to rank the children on speech according to the instructions given in Appendix D. The names of the children to be ranked were listed on a sheet attached to the instructions. Below the names were a series of numbers with a blank space to the right of each number. The aide was instructed to place the name of the child who spoke best on the blank next to number one, the name of the child who spoke next best on the blank next to number two, and so on until all of the children were ranked. From one to two weeks later the aides ranked the children of their respective cottages on non-speech communication according to the instructions given in Appendix E. The same procedure as described above was used to obtain the ranks.

The aide rankings were made from three to five months after the PLS tests had been administered. Only those subjects were included in an analysis who had scored other than zero on the subsection being evaluated.

The data were analyzed separately for each cottage, resulting in five replications of each correlational study. The cottage aide ranks were reversed so that a positive correlation always indicated a

TABLE 1.11. Correlation between PLS vocal scores and cottage aide ranks for speech.

Cottage	*Sex*	*N*	*Mean Vocal Score*	*SD*	*r*
1	F	18	129.4	13.3	.86
2	F	30	166.9	18.2	.71
3	M	22	134.4	16.5	.33
4	M	26	158.0	18.4	.54
5	M	14	169.7	16.6	.64

positive relationship between behavior in the cottage and behavior in the test situation. The product moment correlation was used instead of the Spearman rank order correlation since both methods give essentially equivalent results (*6*).

Table 1.11 shows the correlations between PLS vocal scores and the cottage aide ranks for speech ranging from .33 to .86 with a median correlation of .64. These relatively high correlations indicate that the PLS vocal has some validity for predicting cottage speech communication within a mentally deficient population.

The correlations between PLS non-vocal scores and aide ranks shown in Table 1.12 range from .18 to .80 with a median correlation of .40. The correlations are in the expected direction but the extreme range indicates either that the test score related differently to the children's behavior in the various cottages, or, more likely, that aides in different cottages use different criteria for ranking children.

Of secondary interest was the manner in which PLS vocal scores related to aide ranks for non-speech communication and the relationship between PLS non-vocal scores and the aide ranks for speech communication. The correlations between PLS vocal scores and aide ranks, shown in Table 1.13, which range from −.63 to .76, were not expected and merit investigation. This finding must mean that non-speech communication and PLS vocal scores have different relationships in different cottages or that aides interpret the instructions given in Appendix D in different ways. The latter possibility seems more likely since there did not seem to be any systematic difference in the means or standard deviations between the subjects of the cottages where positive correlations were obtained and those where negative correlations were obtained. It seems likely that certain aides weigh speech behavior

TABLE 1.12. Correlation between PLS non-vocal scores and cottage aide ranks for non-vocal communication.

Cottage	*Sex*	*N*	*Mean Non-Vocal Score*	*SD*	*r*
1	F	21	146.8	22.4	.18
2	F	28	161.6	14.8	.56
3	M	23	149.9	23.7	.80
4	M	27	170.9	11.5	.40
5	M	12	171.7	9.6	.19

TABLE 1.13. Correlation between PLS vocal scores and cottage aide ranks for non-speech communication.

Cottage	*Sex*	*N*	*Mean Vocal Score*	*SD*	*r*
1	F	18	129.4	13.3	—.63
2	F	29	168.9	17.4	.68
3	M	22	133.4	16.5	.49
4	M	26	158.0	18.4	—.46
5	M	13	172.9	12.4	.76

negatively when making judgments concerning the child's non-speech communication while others either weigh it positively or make their judgments of non-speech communication independent of the child's speech skills.

The correlations between the PLS non-vocal scores and the aide ranks for speech, shown in Table 1.14, range from .13 to .52 with a median correlation of .23.

Summary and Conclusions

The discussion of Section One was divided into the presentation of a rationale for language assessment which is related to a general behavioral system, a description of a language test derived from the rationale, and data relevant to the evaluation of the speech and language test and the conceptual system from which it was derived. The classification system used as a guide in developing items to sample language behavior was primarily based on (a) whether the language was vocal or non-vocal and (b) the conditions controlling its occurrence. In line with this rationale, the following seven subtests were developed:

1. *Tact.* In this subtest the examiner presents an object or picture and asks, 'What is it?' The controlling stimulus is the picture or object and the correct response is vocal.

2. *Echoic.* In this subtest the child is asked to repeat digits, words, and sentences. The controlling stimuli are vocal and the response is vocal and bears a point-to-point relation to the stimulus.

3. *Intraverbal.* The examiner asks the child questions such as, 'What do you do when you are hungry?' The stimulus is vocal, the response is vocal, but unlike the response in the previous subtest, it does not bear a point-to-point relation to the vocal stimulus.

4. *Comprehension.* The examiner asks the child to execute a series of commands. The commands are given by speech, by gestures, and by speech and

TABLE 1.14. Correlation between PLS non-vocal scores and cottage aide ranks for speech.

Cottage	*Sex*	*N*	*Mean Non-Vocal Score*	*SD*	*r*
1	F	22	145.0	22.3	.52
2	F	29	161.0	14.9	.52
3	M	23	149.9	23.7	.21
4	M	27	170.9	11.5	.23
5	M	13	169.9	13.3	.13

gestures combined. Thus the controlling stimulus can be either vocal or non-vocal. The correct response is a motor act.

5. *Echoic Gesture.* The examiner demonstrates a series of motor acts which the child repeats. The controlling stimuli are non-vocal; the response is non-vocal and bears a point-to-point relation to the stimulus.

6. *Intraverbal Gesture.* The examiner asks the child a series of questions which can be answered by gestures. The controlling stimulus is vocal; the response scored is non-vocal.

7. *Mand.* The examiner presents the child with a series of situations in which the appropriate response would be to ask a question or make a request. For example, the examiner might ask the child to draw a picture but fail to make paper and pencil available. A correct response could be either a vocal or a non-vocal request.

Three subtests measure vocal behavior, three measure non-vocal behavior, and one, the mand subtest, measures both vocal and non-vocal behavior.

Four adults without previous experience were selected and given intensive training, which included study of the test items, test demonstration, and supervised practice testing. The examiners tested 275 mentally retarded children between the ages of six and 15. Children were randomly assigned to each examiner.

The equivalence of results obtained by different examiners was evaluated by (a) comparing distributions obtained by analysis of variance and the Kolmogorov-Smirnov tests for significance of maximum differences between cumulative proportions and (b) correlating item difficulty levels for pairs of the four examiners. These analyses indicated that the examiners obtained similar results for all of the subtests except the mand subtest. Examiners obtain significantly different results on both the vocal and non-vocal measures of the mand test. In view of this fact, the mand subtest was excluded from most subsequent analyses.

The reliability of the PLS subtests and subsections were evaluated by both odd-even and test-retest reliability procedures. The odd-even reliability coefficients ranged from .84 to .96 on all of the subtests but the mand. In no case was the odd-even reliability for the mand subtest above .40. The test-retest correlations for two samples of 20 subjects drawn from the total population and retested after a two- to five-month period ranged from .86 to .99 with a median coefficient of .92 for the three vocal subtests and from .64 to .92 with a median coefficient of .80 for the non-vocal subtests. The test-retest correlations for 32 children between nine and 12 years of age who were tested after a seven-month period ranged from .82 to .96 on the three vocal subtests and from .59 to .89 on the non-vocal subtests. The test-retest for a group of newly admitted boys who were tested after a 15-month period ranged from .62 to .74 on the vocal subtests and from .15 to .59 on the non-vocal subtests.

The evaluation of the intersubtest relationships showed that the three vocal subtests were highly related with correlations ranging from .77 to .82. Moreover, the correlations between the echoic gesture and comprehension subtests and the three vocal subtests ranged from .63 to .79. These findings indicate that it may not add much to the predictive value of the PLS to maintain

these five subtests as separate categories. However, the correlations between the intraverbal gesture subtest and the others were low, ranging from .01 to .32. The intraverbal gesture subtest would appear to merit more complete study.

Evaluation of the predictive utility of the PLS subsection scores was made in three ways: first, by correlating PLS subsection scores with Wechsler Scale IQs; second, by relating PLS subsection scores with cottage aide ranks for speech and non-speech communication; and third, by classifying subjects into 'high' and 'low' subjects on the PLS and then seeing if they performed differently in a laboratory experiment (see Sections Two and Three). Correlations between the PLS vocal scores and the Wechsler Scale Verbal IQ (with age partialled out) were .68 and .74 for the younger and the older groups, respectively, and correlations between the PLS total score and the Wechsler Full Scale IQ were .41 and .51 with age partialled out. However, the correlation between Wechsler Scale Performance IQs and the PLS non-vocal scores was essentially zero.

The correlations between the PLS vocal score and cottage aide ranks of children for speech communication in five cottages ranged from .33 to .86 with a median correlation of .64. The correlation between PLS non-vocal scores and cottage aide ranks for non-speech communication ranged from .18 to .80 with a median correlation of .40.

In brief, it appears that when the examiner training procedures described are used, the PLS yields reliable data which are relatively unaffected by examiner bias (with the exception of the mand subtest). The vocal and non-vocal subsections seem to be sampling relatively independent behavior. Finally, the PLS subsection scores are usfeul in predicting at least a limited range of non-test language behavior.

Acknowledgment

The writer wishes to thank Dr. Lloyd Dunn for permission to use selected pictures from the plates of the preliminary Peabody Vocabulary Test.

References

1. McCarthy, Dorothea A., Language development in children. Chapter 9 in Carmichael, L., *Manual of Child Psychology*. New York: John Wiley & Sons, 1946, 476-581.
2. Mowrer, O. H., Speech development in the young child: 1. The autism theory of speech development and some clinical applications. *J. Speech Hearing Dis.*, 17, 1952, 17, 263-268.
3. Mowrer, O. H., The psychologist looks at language. *The American Psychologist*, 9, 1954, 660-694.
4. Osgood, C. E., *Method and Theory in Experimental Psychology*. New York: Oxford University Press, 1953.
5. Osgood, C. E., Suci, G. J., and Tannenbaum, P. H., *The Measurement of Meaning*. Urbana: University of Illinois Press, 1957.
6. Peters, C., and Van Voorhis, W., *Statistical Procedures and Their Mathematical Bases*. New York: McGraw-Hill, 1940.
7. Skinner, B. F., *Science and Human Behavior*. New York: Macmillan Co., 1953.
8. Skinner, B. F., *Verbal Behavior*. New York: Appleton-Century-Crofts, Inc., 1957.
9. Walker, Helen M., and Lev, J., *Statistical Inference*. New York: Henry Holt & Co., 1953.

Section Two

Language Behavior of Adults and Retarded Children in Interpersonal Assemblies

Language may be considered a series of verbal interactions between two (or more) persons in which the behavior of each person is at least partially under the control of the other. Thus, any time two persons are assembled, though one be designated 'patient' and the other 'clinician,' or one 'subject' and the other 'experimenter,' the nature of the assembly or grouping itself may affect the behavior of each of the individuals. In interactions between retarded children and teachers or clinicians, for example, the adult is typically considered to be in some manner manipulating or modifying the behavior of the child. However, the child may in turn exert considerable influence over the persons in his environment. In the realm of language the degree of verbal stimulation the child experiences may be related to the kinds of verbal or other cues he presents. Thus, the child whose speech behavior is severely limited or inappropriate may evoke patterns of responses from adults that discourage him from extending or improving his verbal performance.

The studies reported in this section represent an initial attempt to consider the role of interpersonal variables in groups involving retarded children and normal adults. The original paradigm for these studies was developed in an unpublished investigation conducted by Spradlin and Rosenberg (*8*). These authors hypothesized that the types of questions posed by adults in an interview situation with high and low level retardates would vary as a function of the verbal facility of the children being interviewed. In particular, it was hypothesized that the proportion of 'binary' to 'multiple' questions would be greater when adults were interviewing children of low rather than high verbal ability. A binary question was defined as one for which no more than two responses were highly probable (Is your name Sam? Do you prefer ice cream or candy?). Questions were considered multiple if they allowed for a greater range of responses (What is your name? What do you like to eat?). It seemed likely that children of low verbal performance would be more prone to respond to binary than multiple questions and that, consequently, the children would tend to 'condition' the adults to use more of these types of questions. It was assumed that the failure of the child to respond to a particular type of adult questioning behavior would cause that behavior to be extinguished.

On the basis of whether they fell above or below a specified cutoff score derived from the intraverbal and intraverbal gesture subtests of the Parsons Language Sample (see Section One), 96 mentally retarded children at the Parsons State Hospital

and Training Center were dichotomized into 'high' and 'low' verbalizers. The subjects of the study were 16 adults (junior college students), each of whom was assigned to interview six children. Depending on the experimental schedule to which he was assigned, the adult interviewed either six high verbal children, six low verbal children, or a combination of three highs and three lows. The adult was asked to obtain as much information as possible from the children on a variety of topics (see topic sheet in Appendix F). All interviews lasted 20 minutes and were tape recorded. These tape recordings were subsequently replayed by two raters who independently classified all questions as either binary or multiple. The results of the study indicated that low level children elicited a larger proportion of binary questions than did the highs in the initial interviews. These differences were not statistically significant, however.

An examination of the Spradlin and Rosenberg investigation suggested several procedural modifications. The method used for dichotomizing children, for example, did not provide for a large separation between high and low level groups, and factors such as intelligibility or articulateness were not explicitly considered. Further, no determination was made of whether low level children did, in fact, perform differently from the high level children in these sessions or of whether or not the lows responded differentially to binary and multiple questions in the assumed manner.

The three additional investigations reported in this section are extensions and modifications of the interpersonal model discussed in the Spradlin and Rosenberg experiment.

1. Adult Verbal Behavior in 'Play Therapy' Sessions with Retarded Children

GERALD M. SIEGEL

The purpose of this study was to test the hypothesis that the verbal behavior of adult subjects would vary as a function of the linguistic level of children with whom they were assembled in a series of permissive 'play therapy-like' sessions.

Method

Children. Eight children from the Parsons State Hospital and Training Center participated in the investigation. The four 'low' verbalizers scored in the lower 52% on the Parsons Language Sample (PLS). The four 'high' level children scored within the upper 25% on the PLS and, in addition, made six or fewer errors on the Templin-Darley Screening Test of Articulation (*10*). Table 2.1 provides descriptive information concerning the children. The high and low level children are well separated with respect to performance on both the PLS and the articulation measure employed. No attempt was made to control distribution of sex, and boys and girls were unequally represented in the high and low groups. An attempt was made, however, to exclude children with gross physical deviations. Children who were judged to have a gross physical anomaly, who had a peculiarity of posture or gait, or who reportedly were not toilet trained were not included in the study.[1]

Adults. The subjects in the study were two *adults*, a 44-year-old housewife and a 22-year-old male college senior. A set of printed instructions (Appendix G) was given to each adult informing him that he was about to participate in an investigation of the communication of retarded children in permissive free-play situations. The adults were not conversant with the hypothesis being tested or with the fact that the children had been classified as high and low.

The adults were informed that they could use whatever techniques they wished to encourage the children to 'express themselves.' To help stimulate the children they were given a number of simple toys (crayons, paper, small animals, etc.) which they could use at their discretion.

Experimental sessions. Two of the high and two of the low level children

Gerald M. Siegel (Ph.D., State University of Iowa, 1957) is Assistant Professor, Department of Speech, University of Minnesota. He was formerly a Research Associate in the Bureau of Child Research, University of Kansas.

[1]The judges were the author and one other member of the Research Department.

TABLE 2.1. Descriptive data for four high and four low verbal level children.

Measures	*High*	*Low*
Sex	2 males 2 females	4 males 0 females
Mean age	11 yrs. 0 mos.	10 yrs. 8 mos.
Mean weighted score on the PLS	59.90	43.22
Mean number of articulation errors on Templin-Darley 50-item test	3.25	29.67*

*Excluding one child who was untestable.

were randomly assigned to each adult. Adult A (female) was assigned four boys and Adult B (male) two boys and two girls. Experimental sessions lasted 15 minutes and were scheduled weekly for 12 weeks. The adults met individually with each of the children assigned to them in a randomly determined order. The sessions were conducted in experimental rooms equipped for auditory and visual observation. All sessions were tape recorded from an adjoining observation room and were subsequently typed by two secretaries specifically trained for this task.

Preparation of transcripts. The criterion measures employed were all derived from typed transcripts of sessions three through 11.[2] Once the typing had been completed, the transcripts were scored by two 'counters,' who extracted the various criterion measures from the transcripts. As a reliability check, both counters independently scored carbon copies of 10 randomly selected transcripts. The Pearson correlations ranged from .96 to .99. In addition, reliability of the typists was determined by having each typist independently prepare transcripts for a random sample of 10 sessions, both typists using the same taped recordings. These two sets of protocols were then analyzed into the various criterion measures and were again correlated. The correlations for the various measures (described below) ranged from .64 to .99 with an average r of .79. Apparently scoring of transcripts can be done with considerable reliability while preparation of protocols from taped recordings is less accurate (*6*).

Results

Total number of intelligible words spoken by the children in the sessions. Criteria for counting words are presented in Appendix H (these are also prototypes). Although the experimental hypothesis of this investigation concerned the verbal behavior of the adults, a measure of child words was obtained in order to determine the appropriateness of the procedures for designating children as high or low verbalizers. It was hypothesized that children classified as low verbalizers would emit fewer words than those designated as high. To check this hypothesis, the total number

[2]Instructions to typists appear in Appendix H. These instructions are 'prototypes' of those used in the present and the following two studies reported in this section. Although the instructions varied somewhat from investigation to investigation, they were judged sufficiently similar that the instructions included in Appendix H could be considered representative.

TABLE 2.2. Mean values for several measures of verbal behavior as a function of assembling two adults with high versus low level children for nine experimental sessions.

Measures	*High Level*	*Low Level*
Child words	490.61	56.44
Speeches	90.28	37.08
Adult *A* words	1156.67	1235.22
Adult *B* words	650.67	820.72
Adult *A* questions*	22.06	26.00
Adult *B* questions*	21.00	28.94
Adult *A* MLR*	5.58	4.87
Adult *B* MLR*	5.02	5.02

*These values are based on a sample of only 50 responses.

of intelligible words in the nine sessions spoken by the four low level children was contrasted with the number spoken by high level children. As may be seen in Table 2.2, high level children uttered considerably more words than did lows. The Lindquist Type III analysis (*4*) presented in Table 2.3 indicates that this difference is significant.[3] Thus children who differed in terms of their performance on the initial selection procedures also differed in their response to the experimental sessions. These findings are of interest in that they reflect favorably on the selection procedures and also in that they establish that there were indeed two 'treatments' administered to the adults.

TABLE 2.3. Analysis of variance of the number of words spoken by four high and four low level children in nine 'play therapy' sessions with two adults.

Source	*df*	*ms*	*F*
Between	7		
Levels (L)	1	3,393,012	10.84*
Adults (A)	1	154,939	<1
L × A	1	59,398	<1
Error	4	312,737	
Within	64		
Sessions (S)	8	20,100	<1
S × L	8	15,636	<1
S × A	8	29,278	1.17
S × L × A	8	35,034	1.40
Error	32	24,851	
Total	71		

*Significant at 5% level of confidence.

Number of speeches. As defined in this study, the number of speeches is an index of conversational exchange. It is derived by noting the number of times an utterance by one of the participants in the session was followed by or interrupted by an utterance of the second person. To the degree that 'conversation' may be characterized as a series of verbal exchanges between participants, this is a measure of conversation within the sessions. It should be pointed out, however, that this measure is always a product of both the child and the adult since, before one person's speech can end, the second person must begin talking. The measure cannot be interpreted as an index of either adult or child verbal behavior alone.

Considerably more speeches occurred in sessions involving high level children than in those with lows (Table 2.2). This difference was evaluated by a Type III analysis of variance (*4*) and the obtained *F* value was significant (Table 2.4).

Measures of adult verbal behavior. The mean values for the various measures of adult verbal behavior are reported in Table 2.2. Since considerable inter-adult variability was noted, sepa-

[3]In this and all other analyses reported in this section, the 5% level of confidence was selected as the test for significance.

TABLE 2.4. Analysis of variance of the number of speeches by two adults in nine 'play therapy' sessions with two high and two low level children.

Source	*df*	*ms*	*F*
Between	7		
Levels (L)	1	50,933.68	10.07*
Adults (A)	1	369.01	<1
L × A	1	1,540.12	<1
Error	4	5,057.04	
Within	64		
Sessions (S)	8	331.44	
S × L	8	862.71	2.56*
S × A	8	855.98	2.54*
S × L × A	8	555.41	1.65
Error	32	336.88	
Total	71		

*Significant at 5% level of confidence.

rate analyses were computed to determine whether or not the experimental effect might be evident in the behavior of either of the adults considered individually. The values for adult questions and adult mean length of response are based on the first 50 adult responses for each protocol, rather than on the entire session.[4] The responses were designated according to procedures outlined by McCarthy (*5*), Davis (*2*), Templin (*9*), and Siegel (*6*). The instructions used for designating responses and determining whether a response was a question or a statement are presented in Appendix H. The number of intelligible adult words are obtained by the same procedure (discussed earlier) used to determine the frequency of words used by children during the sessions. The number of questions within the 50 responses was obtained by a simple counting procedure. The adult mean length of response (MLR) in this investigation was obtained by dividing the total number of words spoken by the adult in each session by 50 (the number of responses specified), and then by deriving the mean of these values for all sessions with high and with low level children.

When subjected to analysis of variance, none of these measures revealed significant differences in adult verbal behavior in response to children of high versus low verbal level. Although the differences were not significant, the data suggest that the amount of adult verbal behavior is increased but that its complexity is decreased in interactions with low level children.

It is also interesting to note that the MLR values found for the adults in the present study (ranging from 4.87 to 5.58) are considerably lower than the norms provided by Templin (*9*) for normal eight-year-olds (MLR = 7.6). This discrepancy may reflect either some difference between procedures in the present study and those of the Templin study or may indicate that adult verbal behavior is depressed in interpersonal contacts with retarded children.

Discussion and Summary

It was hypothesized that adults would perform differentially when assembled with children classified as 'high' versus 'low' verbal level in a series of 'play therapy' sessions. Tape recordings and typed transcripts were obtained and scored for sessions three through 11 for each of the child-adult groupings. Measures of adult words, adult questions, and adult MLR failed to reveal significant differences in verbal behavior of adults

[4]Actually the first 60 responses were specified; but, according to procedures used by McCarthy (*5*), the first 10 were discarded since these are somewhat unreliable.

when assembled with high or low level children. Children designated as highs did use significantly more words during the sessions than did the lows, indicating that the selection techniques used were predictive of subsequent verbal behavior of the children. Further, a measure of verbal or conversational exchanges, that is, number of 'speeches,' revealed that significantly more of these exchanges occurred when adults were assembled with high rather than low level children.

Adult MLR, while not differing as a function of level of the children, was considerably lower than MLR values reported for eight-year-olds by Templin (*9*). It would be extremely interesting to know whether the low MLR values obtained for adults in the present study are related to the procedures for obtaining and analyzing the speech samples or whether these low values may be attributable to a general depression of adult verbal behavior in interpersonal contacts with retarded children. In this context, studies of adults assembled with *normal* children labeled as retardates or simply brought to an institutional setting for the experimental sessions might prove elucidating.

2. Verbal Behavior of Adults in Two Conditions with Institutionalized Retarded Children

GERALD M. SIEGEL

JEROME P. HARKINS

Studies discussed earlier in this section described an interpersonal model for the verbal interactions between adults and children. Within this model attention is focused on the ways in which each of the participants in an interpersonal setting may influence and modify the behavior of the other. In both studies methodological problems were encountered relating to selection and classification of children, number and type of adult subjects, use of relevant and reliable criterion measures, and specification of the interpersonal situation within which to observe the children and adults.

In the present investigation consideration was given to the effects of assembling adults with mentally retarded children of varying verbal skills and to the particular structure provided for the adult in his interactions with the child. It should be noted that, in terms of the major hypotheses of the study, the subjects are the *adults* rather than the children. Several questions were posed:

1. Will adults respond differentially in their verbal interactions with high versus low level children?

2. Will adults respond differentially in a relatively free and unstructured situation in contrast to a tutorial situation in which they are required to teach the children a specified task?

3. Will adults respond differentially as a function of some interaction between the level of the child and the type of situation?

Method

Selection of adults. Subjects for the experiment were 21 white male students enrolled in the Parsons Junior College who volunteered in response to a request submitted through the office of the Dean of Men. Other than for sex and race, no attempt was made to control selection of subjects within the Junior College. The subjects ranged in age from 18 to 20 years with a mean of approximately 18.5 years.

Selection of children. The children were 42 male residents of the Parsons State Hospital, selected on the basis of age, verbal level, and physical appearance. Initially all boys between seven years, seven months and 15 years, one month were screened for inclusion in the study. On the basis of score on the vocal portion of the PLS, two groups of children were specified, a 'high' group composed of those whose verbal scores

Jerome P. Harkins (M.A. University of Kansas, 1962) is currently at Fordham University.

TABLE 2.5. Comparisons of two groups of mentally retarded boys with regard to age, PLS, and articulation.

Measures	*High Group N = 21*	*Low Group N = 21*	*t*
CA in months			
Range	119-183	91-181	
Mean	162.8	139.3	3.15*
PLS Verbal Scores			
Range	59.3-69.3	45.0-52.0	
Mean	63.49	48.62	16.51*
Templin-Darley Test			
Range	13-49	0-33	
Mean	38.0	6.05	18.87*

*Significant at the 5% level of confidence.

were in the highest 25% and a 'low' group composed of children whose scores fell in the quartile below the median. The children were also administered a version of the Templin-Darley Screening Test of Articulation (*10*). The items within the test were randomized and presented orally without pictorial stimulation. The child's score was the number of correctly articulated sounds. In a previous investigation (7) the examiner obtained a Pearson correlation of .98 when he administered the same test to 22 residents of the institution on two occasions. The mean difference between occasions was slight and non-significant.

Table 2.5 presents a summary of the characteristics of the two groups derived by the above procedures. An examination of this table reveals that the highs were somewhat older than the lows and that both groups had a wide age range. Despite the difference in mean age, there was considerable overlapping between the groups.[5]

Experimental conditions. A high and a low child were randomly assigned to each adult. Two experimental conditions were defined: (a) The 'Unstructured' period in which the adult was left with the child for five minutes and was told that the equipment was being readied and (b) the five-minute 'Structured' or tutorial period which directly followed the Unstructured period and in which the adult was to instruct the child in how to assemble a form-board. Though both the Structured and Unstructured conditions were recorded and subsequently analyzed, the subject was not informed that the initial five-minute period was part of the experiment.

Procedure. When an adult reported for the experiment, he was given a copy of printed instructions (Appendix J) and a diagram of the completed form-board (Figure 2.1). The experiment was described as an investigation of the effects of individual instruction on the learning behavior of retarded children. The adults were also informed that the children would later be tested on the task and that the quality of instruction would be evaluated in terms of the speed and accuracy with which the form-board was assembled.

The experimental sessions were conducted in a sound treated room equipped with a one-way vision mirror and two ceiling microphones which led to a tape recorder in an adjoining observation

[5]Children who were judged deviant in physical appearance, posture, or gait were excluded from the study. A research assistant rated each child on a five-point scale of physical attractiveness, with 5 corresponding to extreme deviation. Those who received ratings of 4 or 5 were excluded except that, in order to obtain the requisite number of subjects, it was necessary to reinstate three of the low level children who had originally received ratings of 4.

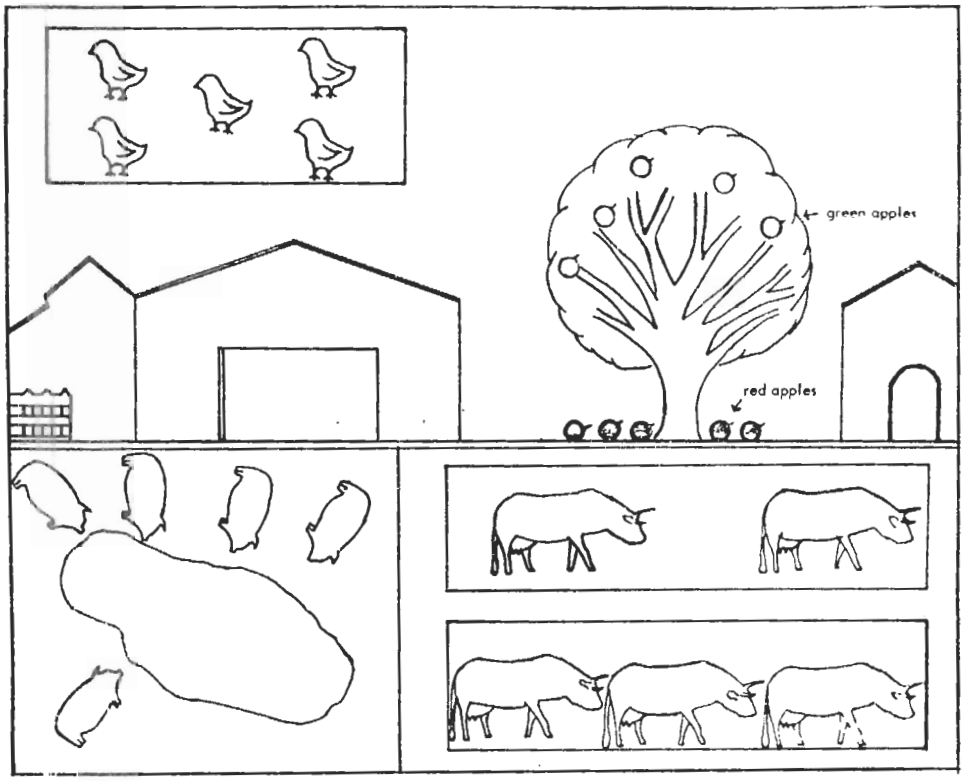


FIGURE 2.1. Diagram of form board in its completed state.

room. The experimental room contained a small table, two chairs, a box which contained the disassembled form-board, and the diagram. Once the adult was settled in the room, the first of his two children (order randomly determined) was brought in and the adult was informed that the experiment would begin in several minutes and that he could use the delay to get acquainted with the youngster. The adult had been preinstructed that a buzzer would sound when the equipment was ready and that the instruction concerning the form-board could then begin. The child was then left with the adult, and the experimenter immediately started timing and recording the session. At the end of five minutes the buzzer sounded and the adult began the formal instruction. Five minutes later the buzzer again sounded and the session was completed. The same procedures were then repeated for the second child.

It should be noted that the design of this study required that the Unstructured condition invariably precede the Structured condition. Thus the effects of order of the conditions were confounded with any differences between the conditions themselves. Such factors as time in the situation or the growth of rapport between the adult and child may have contributed to differences be-

tween the conditions. The influence of these factors cannot be determined by the procedures of this investigation.

Preparation of protocols. When all sessions were completed, the 42 tapes (21 child-adult combinations, two conditions for each) were randomly divided between two trained secretaries. The order in which the Structured and Unstructured conditions on each tape were to be typed was also randomized. The typists were required to transcribe the sessions, indicating where responses began and ended and whether they were statements or questions. In addition, they kept a running count of all responses. Instructions for these tasks are presented in Appendix H. Typist reliability was determined by having each typist independently prepare the transcripts for 10 sessions. The criterion measures to be analyzed were extracted from both sets of protocols by two additional persons who served as 'counters' and the results were correlated. Correlations from the two sets of data for the various criterion measures ranged between .95 and .99. Subsequent self-agreement correlations for the two counters ranged from .97 to .99.

Results

Validity of selection criteria for children. The number of intelligible words spoken by high versus low level children in the experimental sessions was compared in order to test the assumption that highs and lows would present different patterns of verbal behavior. In the Unstructured condition, high level children used an average of 203.57 words in contrast to an average of 92.05 words for the lows. In the Structured condition the highs used a mean of 72.43 words and the lows a mean of 56.28. These data were analysed in a 'Type I' analysis of variance (*4*), and the *F* corresponding to verbal levels of the children was significant (Table 2.6). The original selection criteria thus seem to be indirectly validated. It is also interesting to note that during the Structured condition the amount of child verbalization dropped off significantly for both high and low level children.

TABLE 2.6. Summary of analysis of variance for number of words used by high and low level children in a Structured and an Unstructured condition. N = 42.

Source	*df*	*ms*	*F*
Between	41	4279.65	
Levels (L)	1	84614.00	37.25*
Errors (b)	40	2271.30	
Within	42	16242.38	
Conditions (C)	1	147505.00	12.14*
L × C	1	48480.00	3.99
Error (w)	40	12154.87	
Total	83		

*Significant at the 5% level of confidence.

Number of vocal response units (responses). Values for the various measures of adult behavior in both conditions are summarized in Table 2.7. A vocal response unit closely approximates the 'remark' or 'response' described by Mc-

TABLE 2.7. Mean adult verbal responses to children of high and low verbal level in a Structured and an Unstructured condition.

	CONDITIONS			
	Structured		*Unstructured*	
Measures	*High*	*Low*	*High*	*Low*
Responses	80.00	89.24	88.71	109.00
Questions	43.20	50.20	23.76	28.60
Words	473.00	447.00	603.00	660.00
MLR	5.92	5.15	6.72	6.05
TTR	.45	.41	.41	.36

Carthy (*5*) and Templin (*9*). The specific criteria for designating responses used in the current investigation appear in Appendix H. The measure provides a general indication of total verbal adult activity within the sessions.

As may be seen in Table 2.7, adults made more responses to low than to high level children in both conditions and also made more responses in the Structured than in the Unstructured situation regardless of the level of the children. These differences were evaluated by an ABS analysis of variance (*4*). The *F* values corresponding to levels of the children and to the experimental conditions were both significant (Table 2.8).

TABLE 2.8. Summary of analysis of variance for total number of adult responses in a Structured and an Unstructured condition with children of high and low verbal levels. N = 21.

Source	*df*	*ms*	*F*
Conditions (C)	1	4257	7.75*
Levels (L)	1	4576	15.05*
Adults (A)	20	2073	14.00*
C × L	1	641	4.33
L × A	20	304	2.05
C × A	20	549	3.71
C × L × A	20	148	
Total	83		

*Significant at the 5% level of confidence.

As reflected by this measure, then, adults did modify their behavior according to particular characteristics of the children; and, when asked to 'teach,' the adults increased their verbal output.

Number of adult questions. Questions were differentiated from statements (Appendix H) by the typists as they prepared the transcripts, and the number of adult questions was subsequently determined. These data are summarized in Table 2.7. The ABS analysis of variance (*4*) corresponding to these results is summarized in Table 2.9. Although more questions were posed to low level children, the difference between the means was not significant. Significantly fewer questions were asked during the Structured condition than during the Unstructured condition.

TABLE 2.9. Summary of analysis of variance for number of questions asked by adults in a Structured and an Unstructured condition with children of high and low verbal levels. N = 21.

Source	*df*	*ms*	*F*
Conditions (C)	1	8702.00	53.09*
Levels (L)	1	774.00	2.94
Adults (A)	20	817.40	6.60*
C × L	1	17.00	<1
L × A	20	263.35	3.13
C × A	20	163.90	1.32
C × L × A	20	123.85	
Total	83		

*Significant at the 5% level of confidence.

Number of adult words. In preparing the transcripts, typists omitted responses that contained any unintelligible words.

TABLE 2.10. Summary of analysis of variance for adult words in a Structured and an Unstructured condition with children of high and low verbal levels. N = 21.

Source	*df*	*ms*	*F*
Conditions (C)	1	618000.00	27.71*
Levels (L)	1	5138.00	<1
Adults (A)	20	115541.60	18.15*
C × L	1	36835.00	5.78*
L × A	20	13031.90	2.05
C × A	20	22304.40	3.50
C × L × A	20	6367.10	
Total	83		

*Significant at the 5% level of confidence.

TABLE 2.11. Summary of matched *t* analyses for number of words used by adults in a Structured and an Unstructured condition with children of high and low verbal levels. df = 20.

Comparison	*Mean Differences*	*t*
High vs. Low for Structured Condition	57.00	1
High vs. Low for Unstructured Condition	26.24	1
Structured vs. Unstructured for Highs	129.90	3.26*
Structured vs. Unstructured for Lows	214.00	6.15*

*Significant at the 5% level of confidence.

The difference in mean number of words used by adults with high and low level children (Table 2.7) was not significant (Table 2.10). Significantly more words were used by adults in the Structured condition, however. Since there was a significant interaction between the experimental conditions and the levels of the children, the various combinations of levels and conditions were individually evaluated by the *t*-tests reported in Table 2.11. Neither of the high versus low differences was significant. However, significantly more words were used in the Structured than in the Unstructured condition in assemblies involving children of both verbal levels.

TABLE 2.12. Summary of analysis of variance for adult MLR in a Structured and an Unstructured condition with children of high and low verbal levels. N = 21.

Source	*df*	*ms*	*F*
Conditions (C)	1	151215.00	8.41*
Levels (L)	1	111034.00	10.96*
Adults (A)	20	58634.25	17.49*
C × L	1	449.00	<1
L × A	20	10123.00	3.02
C × A	20	17987.45	5.36*
C × L × A	20	3352.40	
Total	83		

*Significant at the 5% level of confidence.

Adult Mean Length of Response (MLR). Mean length of response has been used extensively as an index of child verbal behavior (*1*, *5*, *6*, *9*). The measure is usually computed for samples of 50 responses. In the present investigation MLR was obtained for the adults and was computed as the total number of intelligible words divided by the number of responses in a given session. In Table 2.3 it may be seen that adults consistently used longer MLR's with high level children, regardless of condidition, and that they had larger MLR values for the Structured than for the Unstructured condition. Both these findings proved significant, as indicated in Table 2.12. Thus, although adults emitted more responses with low than with high level children, the responses were less complex when directed to low level children.

Adult Type-token Ratio (TTR). None of the investigations reported previously in this section used TTR as a measure of adult verbal behavior. It was computed for the middle 200 adult words in each protocol except for two which contained less than 200 words. In these instances (90 and 111 words) the measure was extracted from the entire transcript. TTR was computed by dividing the number of different words (types) by the total number of words (tokens) sampled. If two words differed in any respect (for example, house vs. houses), they were counted as two 'types.' TTR has been interpreted as a measure of redundancy or of 'verbal diversification' (*3*).

TABLE 2.13. Summary of analysis of variance for adult TTR's in two conditions with children of high and low verbal levels. N = 21.

Source	*df*	*ms*	*F*
Conditions (C)	1	398.00	11.42*
Levels (L)	1	398.00	11.83*
Adults (A)	20	33.60	1.22
C × L	1	2.00	<1
L × A	20	33.65	1.23
C × A	20	34.85	1.27
C × L × A	20	27.35	
Total	83		

*Significant at the 5% level of confidence.

In the present investigation TTR values were significantly greater when adults were assembled with high level children than when they were with lows, and in the Unstructured condition than in the Structured condition (Tables 2.7 and 2.13). The fact that adults were redundant (lower TTR) with low level children is consistent with data concerning MLR and adult responses in this investigation, and with the initial hypothesis. Use of lower TTR's in the Structured, or tutorial, condition may have been due to the fact that the adults were required to talk about a limited number of objects, the parts of the form board.

Discussion

The findings of the present investigation are generally supportive of the hypothesis that adult verbal behavior is affected by characteristics of the children with whom they are assembled in two-person situations, as well as being affected by the nature of the assembly (that is, a Structured vs. Unstructured series of interactions).

The obtained differences in verbal behavior, of course, do not establish that a particular pattern of adult behavior will have a 'good' or 'poor' effect on the child's verbal development. While it may be entirely proper for adults to modify their behavior appropriately to the verbal characteristics of the children, it may also be that these modifications deprive the child of the necessity or opportunity for verbal enrichment. Speech pathologists frequently cite lack of adequate speech stimulation and motivation as primary causes of improper or delayed language development. The kinds of responses made by adults in this investigation may be reflective of linguistic deprivation which the retarded child suffered earlier at home.

A comparison of the two experimental conditions reveals that adults did more talking, asked fewer questions, and generally allowed fewer opportunities for the children to respond during the Structured condition. In contrast, the children used fewer words during the Structured than during the Unstructured condition. It would be interesting to determine whether or not instructors of varying levels of effectiveness differ in the amount of verbalization they utilize in an instructional context. It seems plausible that a more effective instructor will structure his interactions with the child so as to encourage the child to verbalize, ask questions, and indicate areas of confusion.

Several additional research directions are suggested by the results of this study. Professional psychologists and special educators are trained specifically to be sensitive to the unique characteristics of the individual child. Is this training reflected in their responses to children of varying linguistic levels? Does such

training in some way create patterns of adult responses that differ from those of untrained persons? Further, the effects of these varying patterns of adult verbal behavior on the *children* should be investigated. What kinds of adult responses are most likely to encourage spontaneous verbalizations from retarded children? How can the adult make speech an attractive alternative for these children? What typical adult responses, on the other hand, tend to depress or discourage language from the children? The next experiment reported in this section represents an initial attempt to study the effects of various patterns of adult verbal response on the language of retarded children.

Summary

The present study was undertaken to test the hypothesis that the verbal behavior of adults varies as a function of the language levels of the mentally retarded children with whom they are assembled. Twenty-one junior college students each met one high level child and one low level child in a Structured (teaching) and an Unstructured (free) situation.

Five measures of verbal behavior of the adults were extracted from highly reliable protocols, which were typed from tape recordings of the sessions. An analysis of the data indicated that with high level children the adults used more responses, greater MLR's, and higher TTR's. Total words and questions did not vary significantly as a function of the levels of the children.

In addition to these results, it was found that the two conditions exerted considerable influence on the verbal behavior of the adults. More questions were asked in the Unstructured condition and more responses were made in the Structured condition. The adults used more words and exhibited higher MLR's and lower TTR's in the Structured condition. Since the Structured condition invariably preceded the Unstructured condition, however, the effects of order were confounded with the effects of the conditions.

3. Verbal Behavior of Retarded Children Assembled with Pre-Instructed Adults

GERALD M. SIEGEL

The studies reported previously in this section have been concerned with modifications in adult verbal behavior that occur when they are assembled with retarded children. The findings show that adults respond differently to high and low verbal level children, but they do not describe the effects of these adult patterns on the verbalizations of the children. From a clinical perspective, then, the question considered in the current study follows naturally: how can adults (clinicians) effectively alter the responses of retarded children in ways that are deemed desirable?

Two modes of adult behavior were selected for study. Observations of adult-child assemblies indicated that adults resort largely to interrogation when they are required to obtain information from these children without benefit of instructions. They barrage the children with questions, changing the subject rapidly, often providing no opportunity for the child to make a response. One experimental 'Interview' condition was devised specifically to represent this mode of adult behavior. By contrast, the author speculated that an effective clinician tends to use few questions when talking to children and that, instead, the clinician attempts to draw the child into verbal exchange rather indirectly, by skillful use of occasional silence, by engaging in verbal play, and by generally reinforcing the child's verbalizations. An attempt was made to create an experimental condition similar in tone to this latter pattern of interaction; this was called the 'Clinical' condition. Descriptions of both conditions appear in Appendixes K and L.

The two conditions were created by providing different sets of pre-experimental instructions to two groups of adults. Effects of the conditions were evaluated in terms of a number of objective measures of child behavior and the effects of the instructions on adult responses were determined by several measures of adult verbal behavior. Finally, since it seemed quite possible that the initial verbal level of the children might be a significant factor in determining their responsiveness to the adults, children of two levels of verbal facility were selected.

Method

Children. The children were 40 girls, ranging in age from 13 years, nine months to 17 years ($M = 15$ years, seven months), who were residential patients at Parsons. Twenty of the children were randomly drawn from the upper 25% of their age range with respect to performance on the vocal portion of the PLS and were designated as 'highs' while the remaining 20 girls had been ran-

domly selected from girls who scored between approximately the 25th and 50th percentiles on the PLS and were designated as 'lows.' No attempt was made in the present study to control systematically for physical appearance or to include articulation performance as an aspect of the child's designated language level. The children were ranrandomly assigned to the two treatments with the restriction that there be an equal number of highs and lows in each.

Adults. The adults were also females and were volunteers from Parsons Junior College. Ten adults were used, ranging in age from 18 years, three months to 20 years, one month, with a mean age of 18 years, 10 months. The adults were selected by the Dean of the College in response to a request from the Research Project. Five of the adults were randomly assigned to the Interview Condition and five to the Clinical Condition. Each adult worked with a total of four children, two highs and two lows. Within a given condition, assignment of children to adults was random.

Orientation sessions. All adults, regardless of treatment, participated in 'orientation' sessions with children similar to those included in the actual experiment. During these sessions each adult saw one high and one low level child and practiced the procedures that had been outlined. These initial contacts with the children also served to familiarize the adults with the institutional setting.

Prior to the first orientation session, each adult was given a set of printed instructions appropriate to the assigned condition, and he followed along as the instructions were read on a tape recorder. The instructions for this initial set of sessions are presented in Appendixes K and L. Adults in the Interview Condition were told that the purpose of the experiment was to determine how well retarded children can be interviewed and what sorts of information can be gleaned from the children. A list of suggested topics (Appendix F) was provided as a guide for talks with the children.

Adults in the Clinical Condition were explicitly cautioned to use only a few questions. They were encouraged to engage in spontaneous verbalizing, without requiring the child to talk; to register approval when the child indicated some desire to talk; and to allow the child's verbal behavior to direct the sessions as much as possible. They were also cautioned to avoid criticizing or punishing the child's speech attempts.

The orientation and experimental sessions were all tape recorded from an adjoining observation facility. The experimental room contained a table, two chairs, and a few small toys. All sessions were timed to 15 minutes.

Experimental sessions. Experimental sessions were held approximately one week after the orientation sessions. The adult met individually with each of the four children assigned her (two lows and two highs) in a randomly determined order. Sessions were again recorded and timed. Just before the first session the adult again read the typed instructions and listened to them on tape. These instructions were essentially a restatement of those of the previous week except that for adults in the Clinical Condition the suggestion was added that silence is sometimes an effective means of encouraging children to verbalize.

Preparation of Transcripts. The tran-

TABLE 2.14. Mean adult verbal responses to children of high and low verbal levels in Clinical and Interview Conditions.

	Clinical Condition			*Interview Condition*		
Measures	*High Level*	*Low Level*	*Combined Levels*	*High Level*	*Low Level*	*Combined Levels*
Responses	190.60	177.30	183.95	200.00	237.80	218.90
Words	1013.20	987.20	1000.20	1127.00	1052.20	1089.60
Questions	61.30	55.70	58.50	149.80	291.90	220.85
MLR	5.42	5.13	5.28	6.01	5.12	5.56
TTR	.472	.470	.471	.442	.384	.412

scripts were typed and scored according to previously described procedures (Appendixes H and I). Though the typists and counters were quite experienced, independent reliability checks were computed for each of the measures extracted from the orientation sessions, and coefficients ranged from .94 to .99. These values are comparable to those found earlier by Siegel and Harkins (Part 2 of this section).

Results[6]

Several measures of adult verbal behavior were computed in order to assess the influence of the instructional procedures on adult performance. Since high and low level children were assembled with the adults in both treatments, it was also possible to obtain a further test of the hypothesis that adults respond differentially as a function of the verbal level of the children. For purposes of analysis, the scores corresponding to the two low level children for each adult were averaged and treated as a single score and scores for the high level children were similarly treated. Data concerning the various measures of adult verbalization, as a function of experimental condition and verbal levels of the children, are presented in Table 2.14.

Number of adult responses. As can be seen in Table 2.14, adults in the Interview Condition made more responses than did those in the Clinical Condition. Further, more adult responses were made in assemblies involving low level children (M = 207.55) than when adults were talking to high level children (M = 195.30). These data were subjected to a Type I analysis of variance (*4*).[7] None of the obtained *F*'s was significant.

Number of adult words. Data concerning the number of words used by adults in the two conditions and with both levels of children are presented in Table 2.14. Once again none of the obtained *F*'s was significant so that the two treatments could not be differentiated on the basis of this measure.

Number of adult questions. Differences in number of questions posed by adults in the two conditions were considerable (Table 2.14) and significant

[6]All of the measures of verbal behavior used in the current investigation have been presented earlier in this section. For fuller descriptions of procedures for obtaining and analyzing the criterion measures presented below, the reader is referred to Parts 1 and 2 of this section.

[7]All of the analyses of variance reported in this investigation are of this type.

TABLE 2.15. Analysis of variance for adult questions in Clinical and Interview Conditions with children of high and low verbal levels.

Source	*df*	*ms*	*F*
Between adults	9		
Conditions (C)	1	252,450.	20.73*
Error (b)	8	12,176.	
Within adults	10		
Verbal Levels (L)	1	6,661.	3.54
CL	1	11,376.	6.04*
Error (w)	8	1,882.	
Total	19		

*Significant at the 5% level of confidence.

(Table 2.15). Although the *F* corresponding to verbal levels of the children was not significant, there was a significant conditions by levels interaction; the appropriate *t*-tests for matched groups were computed and are presented in Table 2.16. These results indicate that for both levels of children the instructions to adults exercised considerable control over their questioning behavior, many more questions being used by adults charged with the specific task of interviewing and accumulating information about the children. Since the two sets of instructions given the adults differed explicitly with regard to use of questions, these findings are quite understandable. The additional finding that adults in the Interview group asked significantly more questions of low than of high level children is of interest in the light of findings in the Spradlin and Rosenberg (*8*) and Siegel and Harkins (Part 2) investigations. The previous authors also found that adults tended to question low level children more than highs, but in neither of the previous studies did the differences reach significance.

TABLE 2.16. Summary of matched *t* analyses for number of questions posed by adults in two experimental conditions with children of two verbal levels.

Comparison	*Mean Difference*	*t*
Clinical vs. Interview for High Level	−88.50	12.46*
Clinical vs. Interview for Low Level	−236.20	9.37*
High vs. Low for Clinical Condition	5.60	2.04*
High vs. Low for Interview Condition	−142.10	15.34*

*Significant at the 5% level of confidence.

Adult Mean Length of Response (MLR). Adults used lower MLR's with low level children than with highs, as had been found previously by Siegel and Harkins. The differences were not significant, however. Nor did the adults differ significantly in MLR usage in the two experimental conditions.

TABLE 2.17. Analysis of variance for adult TTR in two experimental conditions with children of high and low verbal levels.

Source	*df*	*ms*	*F*
Between adults	9		
Conditions (C)	1	168.20	5.32*
Error (b)	8	31.62	
Within adults	10		
Verbal Levels (L)	1	45.00	6.70*
CL	1	39.20	5.83*
Error (w)	8	6.72	
Total	19		

*Significant at the 5% level of confidence.

Adult Type-token Ratio (TTR). TTR values were computed for the middle 200 adult words in each session and are reported in Table 2.14. The analysis of variance summarized in Table 2.17 reveals that all obtained *F*'s are sig-

TABLE 2.18. Mean verbal responses of children of high and low verbal levels in Clinical and Interview Conditions.

	Clinical Condition			*Interview Condition*		
Measures	*High Level*	*Low Level*	*Combined Levels*	*High Level*	*Low Level*	*Combined Levels*
Responses	132.20	68.90	100.55	155.40	115.50	135.45
Words	565.80	106.30	336.05	385.30	110.60	247.95
MLR	6.39	2.16	4.28	4.58	1.81	3.20

nificant, including the conditions by levels interaction. Despite the significant interaction, none of the matched *t*-tests computed for each of the verbal levels and for both conditions was significant. Adults in the Clinical Condition used significantly higher TTR's overall than did adults in the Interview Condition, however. These findings are consistent with those of Siegel and Harkins, who obtained significantly higher adult TTR's when adults participated in a relatively unstructured situation rather than a structured one. The current findings that higher TTR's are used with high than with low verbal level children also accord with those reported by Siegel and Harkins.

Child measures. Verbal behavior of the children in the two conditions was evaluated in terms of number of child responses, child words, and child MLR. Mean values for these analyses (Table 2.18) indicate that children in the Clinical Condition used more words, longer MLR's, but fewer responses. The analyses of variance corresponding to these measures did not reveal significant differences as a function of either experimental condition or the conditions by levels interactions. For all measures, as was expected, children designated as high level responded with significantly longer or more complex responses than did the lows.

Discussion

Analyses of adult behavior were undertaken because responses of the adults to the Control and to the Interview instructions were crucial in establishing the experimental conditions. Of the five measures of verbal behavior obtained, only number of questions and TTR revealed significant differences between conditions. Adults instructed in Interview procedures asked many more questions and used smaller TTR's than did those in the Clinical group.

Despite the suggestion in the Clinical instructions that silence may be a useful technique for encouraging the children to talk, the two verbal output measures (number of responses and words) which should be related to the use of silence did not yield significant differences between the two groups of adults. These results may indicate that while the *type of situation* was instrumental in influencing adult verbal behavior, the *pre-experimental instructions* were not adequate for this purpose.

The inclusion of both high and low level children in the present investigation permitted another evaluation of

the hypothesis that linguistic level of the children exerts some control over adult verbal behavior. Adults used significantly more questions and smaller TTR's with low level children. Although results concerning number of responses were not significant, they were in the same direction (more responses to lows) as the significant findings in the Siegel and Harkins investigation (Part 2). The TTR analyses are particularly noteworthy in that this measure has consistently revealed differences in adult response to high and low level children.[8]

Despite the finding that adult responses (TTR and number of questions) were subject to the influence of both experimental condition and level of the children, none of the measures of child performance differentiated children in the Clinical and the Interview Conditions. Apparently, within the situations investigated, variations in patterns of adult behavior are not necessarily followed by corresponding variations in the verbal behavior of children. In future investigations it may be useful to provide more extensive instruction for the adults and to arrange for more numerous assemblies rather than the single 15-minute experience provided in the current study.

Summary

It was hypothesized that adults trained in a relatively permissive and 'clinical' approach would evoke more verbal responses from retarded children than would adults who were simply required to interview the children. Five adults were assigned to the Clinical and five to the Interview Conditions. Each adult met individually with four children, two of 'high' and two of 'low' verbal level for 15-minute sessions. All sessions were recorded, transcribed, and subsequently analyzed in terms of a number of objective measures of adult and child verbal behavior. Of five measures of adult behavior employed, on only two (number of adult questions and TTR) did adults in the Clinical Condition differ from those assigned to the Interview Condition. Adults responded differentially to the verbal behavior of high and low level children. This finding supported the results of a previous investigation. None of the measures of child verbalization revealed any differences between the children assigned to the two conditions.

[8]In the light of these findings, TTR values were computed for a replication of the Spradlin and Rosenberg (*8*) investigation cited earlier in this section. Although none of the other measures of adult behavior differentiated assemblies with high versus low level children, TTR values were significant. Once again adults used smaller TTR's with low level children.

References

1. Darley, F. L., and Moll, K. L., Reliability of language measures and size of language sample. *J. Speech Hearing Res.*, 3, 1960, 166-173.
2. Davis, Edith A., *The Development of Linguistic Skill in Twins, Singletons with Siblings, and Only Children from Age Five to Ten Years.* Institute of Child Welfare Monograph No. 14. Minneapolis: Univ. Minn. Press, 1937.
3. Jaffe, J., Language of the dyad: a method of interaction analysis in psychiatric interviews. *Psychiatry*, 21, 1958, 249-258.
4. Lindquist, E. F., *Design and Analysis of Experiments in Psychology and Education.* Cambridge: The Riverside Press, 1953.
5. McCarthy, Dorothea A., *The Language Development of the Preschool Child.* Institute of Child Welfare Monograph No. 4. Minneapolis: Univ. Minn. Press, 1930.
6. Siegel, G. M., Inter-examiner reliability for mean length of response. *J. Speech Hearing Res.*, 5, 1962, 91-95.

7. SIEGEL, G. M., Experienced and inexperienced articulation examiners. *J. Speech Hearing Dis.*, 27, 1962, 28-35.
8. SPRADLIN, J. E., and ROSENBERG, S., Complexity of adult verbal behavior in a dyadic situation with retarded children. Parsons Project Working Paper 18, Nov., 1959.
9. TEMPLIN, MILDRED C., *Certain Language Skills in Children.* Institute of Child Welfare Monograph No. 26. Minneapolis: Univ. Minn. Press, 1957.
10. TEMPLIN, MILDRED C., and DARLEY, F. L., *The Templin-Darley Tests of Articulation.* Iowa City: Univ. of Iowa Bureau of Educational Research and Science, Extension Division, 1960.

Section Three

Effects of Consequences on Vocal Behavior

How do events which follow vocal responses affect subsequent vocalizations? This question is pertinent both to the study of complex processes of communication in naturalistic settings and to the systematic investigation of language learning in abstracted experimental settings. Yet research evidence about this problem is in short supply.

Speech therapy experiences raise many questions about the conditions under which changes in language behavior occur. For instance, how frequently should the clinician 'reward' the child for a correct response? What kinds of rewards are likely to bring about the most rapid and permanent learning? What happens when a child hears his own voice at different intervals of delay? Answers to these questions are not readily available to the clinician because it is difficult to generalize from the few available experimental studies. Yet it is possible that research will eventually offer such answers for use in complex therapeutic situations.

The two studies which follow investigate the effect of consequent events on vocal behavior. The first by Horowitz studied the effect of some reinforcing variables on the learning of vocal responses. The second study by Copeland was concerned with the use of feedback modification to facilitate vocal productions. Both studies used the Parsons Language Sample as a main criterion measure for selecting subjects, and both emphasized the use of observable events in studying language behavior.

1. Partial and Continuous Reinforcement of Vocal Responses Using Candy, Vocal, and Smiling Reinforcers among Retardates

FRANCES DEGEN HOROWITZ

Early childhood vocalizations are made up of a variety of individual sounds. These are closely followed by the development of discriminable responses, intelligible speech, and finally the utterance of complex and abstract concepts. Though differing in complexity and meaning, all vocalizations can be thought of as a class of responses. Some utterances are likely to be followed by a variety of events or consequences. Some of these consequences lead to changes in the frequency of particular kinds of vocalizations. Such consequences are called reinforcement. In other words, reinforcement as used in this paper refers to any event or consequence which changes the frequency of the response it follows. Some consequences may lead to a decrease in particular kinds of vocalizations and may be designated as punishing events or stimuli. Other consequences may lead to an increase in vocalizations and may be designated as *positive* reinforcers.

This study was concerned only with those events which were followed by an increase in vocalizations. However, if the process of language learning is to be fully understood, we will probably need to know more about a wide range of reinforcing events and also the amount of reinforcement needed for certain responses or response patterns to develop. Conversely, such information may also point to the process by which inadequate language development occurs.

One class of reinforcing events thought to be relevant to vocalizations is the action of another person. For example, a child's vocalizations are often followed by adult (as well as peer and sibling) responses such as a smile, laughter, verbal output, or physical contact. If these responses or personal reactions have the effect of increasing the frequency of the vocalizations which they follow, they would be designated as reinforcing events or stimuli. Objects constitute another class of reinforcers. For instance, if vocal requests such as 'I want a drink of water,' 'Give me some peanut butter and crackers,' 'Let me have that truck' increase in frequency when they are followed by presentation of the object, then the objects are reinforcers.

Frances Degen Horowitz (Ph.D., State University of Iowa, 1959) is Research Associate in the Bureau of Child Research and Assistant Professor in Child Development, University of Kansas. She is currently a Post-Doctoral Fellow in the Bureau. She wishes to acknowledge the critical help of Drs. Joseph E. Spradlin and Gerald M. Siegel in the conduct of this study. To Mrs. Eleanor Kobel and Mr. Bob Brandenburg is due appreciation for their interest and care in executing their roles as experimenters.

Although an object may be received at the hands of another person, it is the receipt of the object itself which is the primary consequence of the vocalization. Thus it is assumed that the object, rather than the other person's actions, takes on the role of the primary reinforcer.

Such descriptive analysis has led to the arbitrary grouping of two classes of reinforcing events: *Social reinforcers*, defined primarily in terms of the behaviors or personal reactions of another person; and *object reinforcers*, defined primarily in terms of the objects received. In the complex of everyday behavior these two kinds of reinforcing events are often mixed.[1]

In laboratory studies of behavior the most commonly used reinforcers have been objects. A motor task or game correctly learned or played leads to a prize of a toy, candy, or gum. This is true for many of the studies concerned with verbal learning in children (see Spiker, *12*). Such reinforcements are easy to describe and define. Social reinforcers, while used occasionally, are usually less adequately defined. Stevenson and his colleagues (*14*, *15*, *16*, *18*) have reinforced motor responses using such social reinforcers as head nods, smiles, and verbal statements about the child's performance. Rheingold, Gewirtz, and Ross (*10*) reinforced infant vocalizations with verbal output, facial expressions, and physical contact. Krasner (7) reports 38 adult verbal conditioning studies in which reinforcement was characterized by experimenter approval indicated by vocalization or body movement. His review indicates that such broadly defined social reinforcers are effective in increasing the frequency of particular classes of responses.

Various types of behaviors have been defined as socially reinforcing: vocalization, facial expression, physical response. In order to delineate the social reinforcement concept further, its component parts should be separately investigated. The present study has attempted to do this.

The percentage of vocalizations of an individual that are followed by reinforcing events is unknown, but it is quite unlikely that reinforcement occurs 100% of the time. Certain arbitrary percentages have been selected for study in order to determine what differences in performance are to be found when percentage of reinforcement differs. Percentages of reinforcement have been employed ranging from zero (no reinforcement) through 20%, 33%, 50%, 67%, 75%, and 100%. Any percentage of reinforcement other than 0% and 100% is referred to as *partial reinforcement* or intermittent reinforcement.

One of the more stable laws of psychology has been the partial reinforcement effect, referred to as PRE (*8*). Empirically it has been noted that partially reinforced responses (particularly when reinforced at least 50% of the time) approach the same level of performance as responses reinforced 100% of the time (*5*, *8*). That is, for responses under partial reinforcement the eventual level of acquisition approaches that of continuously reinforced responses. Some recent

[1] *Object reinforcement* can actually be thought of as a form of tertiary reinforcement, while *social reinforcement* is a type of secondary reinforcement. There are also reinforcers which fall into neither category. A discussion of these matters, however, is beyond the scope of this paper. Here primary reinforcer is used in its non-technical sense to refer to the most potent reinforcer in a given situation.

evidence suggests that if partial reinforcement is continued over enough trials, the level of acquisition may be higher than for continuously reinforced responses (*10*). It has also been found that when reinforcement ceases during extinction, those responses which were partially reinforced persist for longer periods of time than those responses which were continuously reinforced. Since persistence of a response during the extinction period is a measure of its resistance to extinction, partially reinforced responses are said to have *greater resistance to extinction* than continuously reinforced responses. The theoretical aspects of PRE have been discussed by Spence (*11*) and Amsel (*1*). The PRE has been demonstrated with many kinds of responses and with many different species. There is some evidence that continuous or very high percentage of reinforcement may be necessary in the early stages of learning (*11*) though this has not been extensively investigated. With normal adult human subjects PRE has been shown to exist in verbal behavior as well as motor behavior (*6*). Two studies of partial reinforcement of motor responses of retarded children have demonstrated the PRE (*3*, *13*). However, not much is known about partial reinforcement of verbal behavior with mentally retarded subjects.

There has been some attention to the effects of social reinforcement on the behavior of mental retardates. Mentally retarded children are particularly responsive to social reinforcement when such reinforcement involves verbal responses and facial expressions on the part of an adult experimenter (*14*, *15*, *16*). These investigators suggested that social reinforcement may be more effective with institutionalized than with non-institutionalized populations because institutionalization represents a form of social deprivation. However, in a later critical formulation Zigler (*17*) indicated that both the preinstitutionalization history of the child and the characteristics of the institution may have important effects upon the efficacy of the social reinforcers. The research of Stevenson and Zigler has involved motor responses. Barnett, Pryer, and Ellis (*2*) socially reinforced verbal responses of mentally retarded children and adults (MA's ranging from seven to 12 years). They found a significant increase in sentences starting with 'I' or 'we' when their occurrence was followed by an experimenter's verbal statement of 'good.' There are, however, few other studies concerned with the effect of types and percentage of reinforcement on vocal responses of mentally retarded children.

The present study was designed to answer the following questions:

1. What types of reinforcing stimuli are most effective in increasing the frequency of a correct vocal response?

2. Can vocal and non-vocal components of social reinforcement be distinguished as two distinct types of reinforcing events?

3. Can a partial reinforcement effect be demonstrated with the vocal behavior of retardates, comparing 50 and 100% reinforcement, and does the percentage of reinforcement interact with type of reinforcement?

Method

Subjects. Subjects were randomly assigned to reinforcement type and reinforcement percentage groups from a

TABLE 3.1. Mean age (in years) of subjects in each experimental group.

Groups	*C* *I*	*V* *II*	*S* *III*	*C-V* *IV*	*V-S* *V*	*Range*
50%	14.50	15.00	14.83	14.50	15.00	12-16
100%	14.16	14.00	14.33	13.33	15.00	11-16
Total	14.33	14.50	14.29	14.08	15.00	11-16

pool of 72 children between the ages of 11 and 16, having Parsons Language Sample (PLS) vocal T scores between 53 and 70.[2] Due to apparatus difficulties, some of the original children were dropped and others substituted. Substitutions were made by randomly drawing from the remaining 12 children. Tables 3.1 and 3.2 indicate the mean age and PLS vocal scores for the 60 subjects whose data were included in analysis of results. (Reinforcement type and percentage level groups will be described below). For 70% of the subjects Wechsler-Bellevue (WB) IQ scores were available. The mean WB verbal scores was 58.76, with a range from 44 to 75; the mean WB composite score was 55.90 with a range from 36 to 78. There were 32 males and 28 females, but groups were not counterbalanced for sex. In all cells, however, at least two out of the six subjects were female.

Most of the children were familiar with the research division of the Training Center and had previously received some kind of prize or candy for participating in research. In general, subjects liked being asked to participate in the 'games.' The experimenters used in the present experiment were unfamiliar to the subjects.

Apparatus. Figure 3.1 represents the apparatus as viewed by the subject and its relationship to the observation booth. The front panel measured approximately 6½ feet × 4½ feet. The sliding panel was at approximately face height of a seated subject and slid sideways. Directly below the sliding panel was a drawer which could be pushed out toward the subject. Both the sliding panel and the reward drawer were operated by the experimenter behind the panel. A throat microphone which triggered a Dual Channel Voice Operated Relay (VOR),

[2]The T scores were based on an earlier analysis of PLS data than that reported in Section One. The T conversion tables are given in Appendix C.

TABLE 3.2. Mean PLS vocal T scores of subjects in each experimental group.

	Treatment Groups				
Groups	*C* *I*	*V* *II*	*S* *III*	*C-V* *IV*	*V-S* *V*
50%	65.22	59.66	51.32	64.03	59.58
100%	61.86	63.66	60.06	59.02	62.03
Total	63.54	61.63	55.69	61.54	60.80

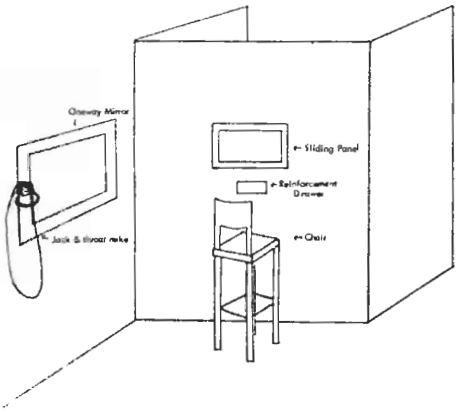


FIGURE 3.1. Diagram of front view of experimental apparatus.

made by the Lafayette Instrument Company, was plugged into a jack on the side of the observation window. This connection led into a standard timer in the observation booth. On the back of the panel were two ordinary doorbell buzzers, which provided the auditory discriminative stimulus for each trial.[3] One was operated manually by the experimenter behind the panel, and one was operable by button press from the observation booth. In the observation booth a button press initiated a buzzer sound for .20 seconds, with automatic offset. The onset of the buzzer also initiated a standard electric timer. The timer was offset by activation of the VOR. That is, when the child made a vocal response, impulses delivered by the VOR offset the standard electric timer. Manual resetting was required. If no impulses were picked up by the VOR in 15 seconds, the timer automatically offset and was reset manually for the next trial. The VOR was adjusted to noise level of the experimental room at a dial level reading of about 6. This setting was sensitive enough to allow operation of the VOR through voice vibrations but was not operated by apparatus noises such as the buzzer and panel operations.

Experimenters and instructions. Two experimenters were involved in the study, E_1 and E_2. E_1, who remained behind the panel at all times, delivered the auditory discriminative stimulus (the buzzer) during the instructional period and the reinforcement at appropriate times. The role of E_2 was to introduce and instruct the child, to initiate the auditory discriminative stimulus for each trial during the experimental session, to record the response and the latency of response, and to reset the timer. At the end of an experimental session E_2 reentered the experimental room and concluded the session.

Both Es went through a training period in which procedures were practiced, and both participated in running pilot subjects where procedural problems were worked out. Neither E was told of the major interests of the study. E_2 was a woman in her middle thirties who had previous experience in working with children. E_1 was a 21-year-old male college student (business administration major) who was working as a summer research assistant.

Subjects were taken into the experimental room by E_2 and given the following instructions:

'Hello. We have a game for you to play—here in this room. This game is fun. Just sit down in the chair.' E_2 then showed the child three cards: a picture

[3] In pilot work a light was placed directly above the sliding panel for use as a visual discriminative stimulus. This proved unsatisfactory since subjects did not always respond to the light's onset and offset. In order to solve this problem of visual attention an auditory discriminative stimulus was substituted.

of a dog, a picture of a cat, and a picture of a bird. 'What is this?' E_2 waited for response to each one. In all cases the children correctly identified the stimuli. 'Now, this is part of the game.' E_2 took throat mike from hook. 'It goes around the neck like this.' E_2 showed the child how it fit on E's neck, let the child feel the mike, and then fastened it to the child's neck. 'When the buzzer rings,' E_1 rang the buzzer, 'you name one of these animals: bird, cat, or dog.' These were given in random order. 'One of these is the right answer.' E_1 rang the buzzer and E_2 waited for the child to respond. For any response E_2 then said: 'That's right.' E_1 rang buzzer and E_2 said, 'Remember, when the buzzer rings, you say one of the words. Just one of them is the right answer.' Then, for groups C and CV (explained below) E_2 said, 'Here is a bag for the things you'll get. When you want to stop you just say stop. Do you have any questions? I'm going to turn the game on now. Remember, when the buzzer rings you say cat, bird, or dog. Just one of them is the right answer.' E_2 then left the room, initiated the buzzer for the first trial, recorded the subject's vocal response, the latency of response, and whether or not the reinforcement was delivered. If the subject did not respond, E_2 reentered the room and repeated the instructions. If the subject still did not respond, he was eliminated from the study. One subject was eliminated from the study for this reason.

Treatment groups and procedure. Subjects were randomly assigned to one of five type of reinforcement groups and each of these groups was then randomly divided into two percentage of reinforcement groups (50% or 100%). A correct response was either cat, dog, or bird, depending upon its random assignment to each subject. The reinforcement groups and the procedures which distinguished them during acquisition were as follows:

Group C: For each rewarded response the subjects in Group C received a packet of five small hard candies. The subjects put them in the bag given them by E_2 during the instructional period.

Group V: For each rewarded response the subjects in Group V were given vocal reinforcement. From behind the panel E_1 said ,'That's right,' or 'Very good,' or 'Fine,' etc.

Group S: For each rewarded response the subjects were given smiling reinforcement. The sliding panel slid back and E_1 smiled broadly, nodded his head, and closed the panel.

Group CV: For each rewarded response the subjects received candy and vocal reinforcement, as described above.

Group VS: For each rewarded response the subjects received vocal and smiling reinforcement, as described above.

All subjects were rewarded for any response during the first five trials of the acquisition phase of training. This reward was found to be necessary during pilot work in order to get further responses. Thereafter, subjects who were under continuous reinforcement received the appropriate reward for every correct response made. After the first five trials subjects who were under partial reinforcement received the appropriate reward for every other correct response. That is, partial reinforcement subjects were rewarded for 50% of their correct responses.

The acquisition phase was concluded when the subject had made five consecu-

tive correct responses or at the conclusion of 30 trials. Then without a break the extinction period began and continued until the subject made four consecutive incorrect responses, indicated he wanted to stop, or had concluded 48 trials. Only one subject requested that a session be concluded, and this was during extinction training.

When extinction was concluded, E_2 reentered the room, removed the throat mike, and ended the session. All subjects who did not receive candy during the training session, that is, subjects in groups V, S, and VS, received a small bag of candy at the end of the session.

Results

The data from this study provided three different kinds of response measures: trials, correct responses, and latency of response.

Number of trials. Analyses of each measure are reported separately. Since the conclusion of both acquisition and extinction depended upon the performance of the subject, it was possible to investigate whether the number of trials in acquisition and extinction differed as a function of type of reinforcement or percentage of reinforcement. Performance in acquisition and extinction was evaluated separately.

Acquisition. Acquisition criterion was reached when five consecutive correct responses were made or when 80 trials had been given. Figure 3.2 indicates the mean number of trials to criterion in acquisition. It can be seen that no subject in the 50% group reached the criterion of five consecutive correct responses; that is, all subjects received 80 trials.[4] No analyses of the number of trials to criteria were performed for the 50% group. In the 100% group 13 of the 30 subjects required 80 trials. Thus all the subjects at the 50% level failed to give five consecutive correct responses, and 43% of the subjects at the 100% level did not give five consecutive correct responses. An analysis of variance (simple randomized design) was carried out on the data for the 100% group, but the F of 1.94 was non-significant. In other words, no differences were found between reinforcement type groups in the number of trials required to reach criterion during acquisition at the 100% level.

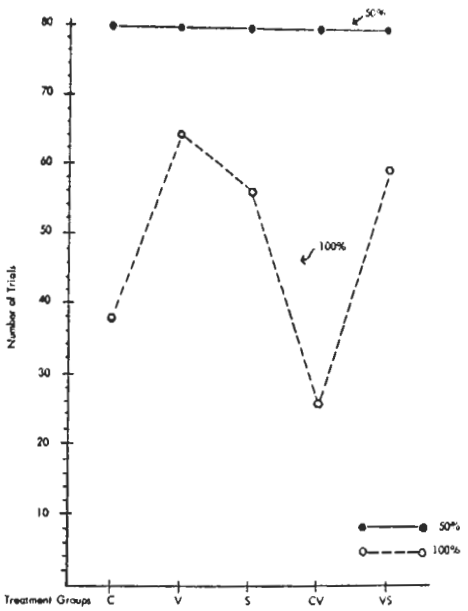


FIGURE 3.2. Mean number of trials to criterion in acquisition for five types and two percentages of reinforcement.

[4]Number of trials *to* criterion includes the number of trials up to but not including the trials on which five consecutive correct responses were made. Thus, if a subject received 80 trials, he did not make five consecutive correct responses.

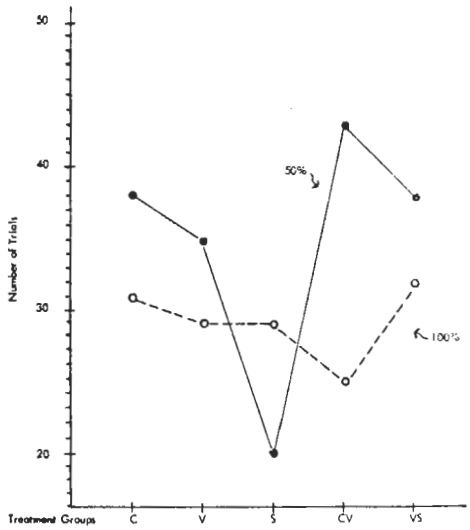


FIGURE 3.3. Mean number of trials to criterion in extinction for five types and two percentages of reinforcement.

Extinction. Extinction criterion was reached when four consecutive incorrect responses were made or when 48 trials had been given. Figure 3.3 indicates the mean number of trials to criterion. The results of the analysis of variance of trials to extinction are shown in Table 3.3. The significant type of percentage interaction was further analyzed by employing a simple randomized design at each percentage of reinforcement. The type of reinforcement effects at each percentage of reinforcement were non-significant. It can be seen from Table 3.4 and Figure 3.3 that the probable source of the significant interaction was in the very few trials required to extinguish Group S at the 50% level. Unlike all other groups at the 50% level, Group S required fewer trials to extinguish than groups at the 100% level.

Correct responses. Within the design of this study it was possible for subjects who learned most slowly to make the most correct responses during acquisition. That is, because the acquisition criterion was five consecutive correct responses, a subject who learned very rapidly could reach criterion on the first five trials. A less alert subject might make 30 correct responses in 80 trials but not make five consecutively correct responses. Thus he would have more correct responses but would not have met learning criterion. An analysis was made to evaluate a possible interaction effect involving type of reinforcement, percentage of reinforcement, and the training phase (acquisition-extinction).

In this analysis, the total number of correct responses was averaged over subjects for each percentage of reinforcement group within each type of reinforcement group for acquisition and extinction. Thus, two between factors

TABLE 3.3. Analysis of variance for trials to extinction.

Source	*df*	*ss*	*F*
Treatment	4	1022.73	1.06
% Reinforcement	1	437.40	1.81
T × % Reinforcement	4	2628.40	2.72*
Within	50	12076.20	
Total	59	16164.73	

*Significant at .05 level of confidence.

TABLE 3.4. Mean trials to extinction.

	Treatment Groups				
Groups	*C*	*V*	*S*	*CV*	*VS*
100%	31.80	29.00	29.07	25.50	32.17
50%	38.33	35.50	19.33	43.33	38.17
Total	35.06	32.25	24.20	34.42	35.17

(per cent reinforcement and type of reinforcement groups) and one within factor (acquisition-extinction) were included in a Type III analysis of variance (*9*). The summary table of this analysis is given in Table 3.5. The main effects of reinforcement percentage and acquisition-extinction were significant. Two interactions were significant: reinforcement percentage × acquisition-extinction, and type of reinforcement × acquisition-extinction. Each of these interactions was further analyzed. The reinforcement percentage × acquisition-extinction analysis indicated that the means for acquisition and extinction at the 50% level were significantly different while the means at the 100% level were not. Thus subjects at the 50% level made significantly more correct responses during acquisition than during extinction. The means in Table 3.6 indicate that subjects at the 50% level made more than twice as many correct responses in acquisition than in extinction, whereas this is not true for subjects

TABLE 3.6. Mean number of correct responses.

Groups	*Acquisition*	*Extinction*
50%	29.33	13.50
100%	21.37	13.43

TABLE 3.5. Number of correct responses, acquisition and extinction, Type III analysis of variance.

Source	*df*	*ss*	*F*
Between	59	3729.49	
% Reinforcement	1	484.01	8.44**
Treatment	4	52.80	<1
% Reinforcement and Treatment	4	327.10	1.43
err (b)	50	2865.58	
Within	60	7565.50	
Acquisition-Extinction	1	4236.41	95.41***
A-E × % Reinforcement	1	468.07	10.54***
A-E × Treatment	4	468.36	2.66*
A-E × % × Treatment	4	172.74	<1
err (w)	50	2219.92	
Total	119	11294.99	

*Significant at .05 level of confidence.
**Significant at .01 level of confidence.
***Significant at .001 level of confidence.

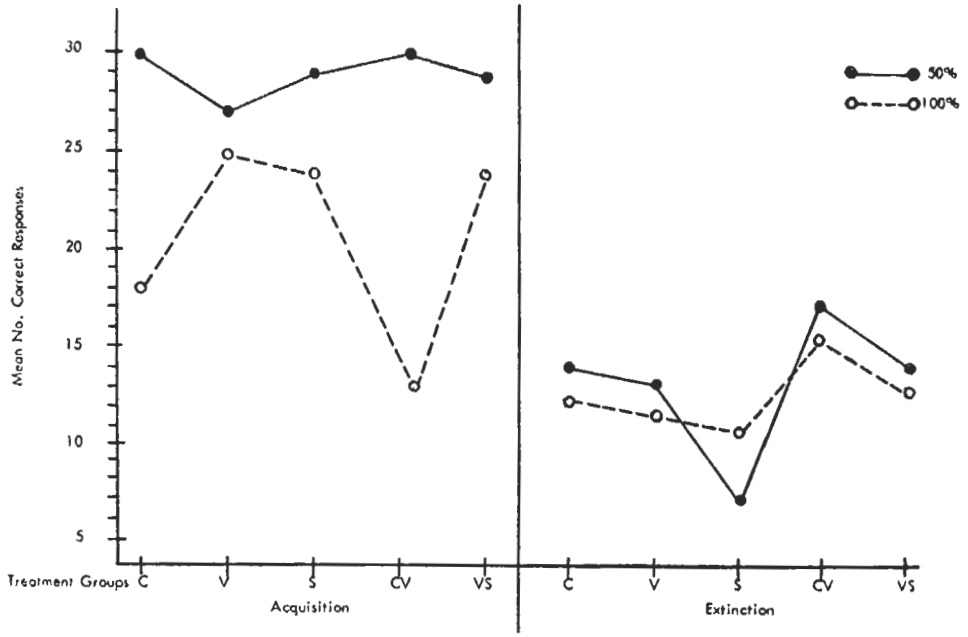


FIGURE 3.4. Mean number of correct responses during acquisition and extinction for five types and two percentages of reinforcement.

at the 100% level. As can be seen in Figure 3.4, the number of correct responses in extinction is practically the same for the two percentages of reinforcement. However, 50% level subjects received about 15 reinforcements whereas subjects at the 100% level received about 21 reinforcements during acquisition.

The analysis of the type of reinforcement × acquisition-extinction interaction indicated that type of reinforcement groups were not significantly different during acquisition but were significantly different in extinction. The *t*-test comparisons of the means shown in Table 3.7 indicate that in extinction Groups C and V made significantly more correct responses than Group S and significantly fewer correct responses than Group CV; Group S made significantly fewer correct responses than Groups CV and VS; group CV made significantly more correct responses than Group VS. Thus, during extinction Group S made the fewest correct responses and Group V made the most. Group S was thus least resistant to extinction and Group CV was most resistant to extinction.

TABLE 3.7. Mean number of correct responses.

Groups	*Acquisition*	*Extinction*
C	24.33	14.00
V	26.50	12.83
S	26.83	9.75
CV	22.25	17.00
VS	26.83	13.75

Latency of Response. A standard electric timer recorded how long it took the subject to respond after the presentation of the buzzer stimulus. This measure constituted the latency of response. Three analyses of this measure were carried out: overall mean latency of response, mean latency of correct responses, and mean latency of incorrect responses. In the last two analyses blocks of trials was included as a factor.

Mean Latency of Response. An analysis of variance of the mean latency of response included the between factors of type of reinforcement group and percentage of reinforcement and the within factors of acquisition-extinction and correct-incorrect response. An extension of a Type VI design (*9*) indicated that the quadruple interaction involving all four factors was significant, but all further analyses proved to be non-significant. One main effect, correct-incorrect response, was significant. The mean latency of correct responses was 1.19 seconds and the mean latency of incorrect responses was 1.44 seconds.

General and Specific Findings. The results of this study may be summarized in terms of two general findings: (a) Partial reinforcement did not lead to greater resistance to extinction but did markedly affect performance during acquisition, and (b) the separate and combined reinforcement conditions produced no apparent effect on behavior during acquisition but did create some differences in extinction behavior. The specific findings were as follows:

1. A combination of candy and vocal reinforcement yielded the greatest resistance to extinction while smiling reinforcement yielded the least resistance to extinction.
2. Candy, vocal, and vocal smiling reinforcers did not differ from each other but yielded greater resistance to extinction than did smiling reinforcement and less resistance to extinction than did candy-vocal reinforcement.
3. No subject who received partial reinforcement reached the acquisition criterion, while 57% of the subjects who received continuous reinforcement reached acquisition criterion.
4. There was no significant difference between partial and continuous reinforcement with regard to extinction.
5. There were no differences in acquisition as a function of type of reinforce administered.

Discussion

In this investigation vocal behavior was studied as a function of type of reinforcement and percentage of reinforcement. Since type and percentage generally did not interact, discussion of the results will focus mainly on the effect of each of these factors separately.

Type of reinforcement. A combination of candy and vocal reinforcement was the most effective reward used in the study. Neither smiling alone nor smiling and vocal reinforcement (the most usual combination used in social reinforcement conditions) was as effective as the combination of candy and vocal reward. These results create some speculation. Vocal reinforcement for a vocal reponse may have provided important feedback stimulation of a type which matched, in kind, the subject's response. While this is not an echoic situation, it may be that vocal reinforcement for a vocal response may create a reinforcement-in-kind condition. We know little of the importance of this kind of reinforcement variable. Now, with candy added to the vocal rein-

forcement condition, vocal reinforcement is enhanced. It cannot be argued that children in this institution were abnormally deprived of candy. Most children's trips to the Research Unit were concluded with a packet of candy or bubble gum. Selections from the canteen stock of candy were available to all children in the institution in exchange for tokens. While candy consumption was controlled, candy was relatively accessible. However, the controlled dispensation of candy probably enhances its value for all children, including the subjects in this study. It is also possible that candy constituted a more 'immediate' reinforcer than did vocal and smiling reinforcement. This may have been particularly the case if subjects who did not receive candy as reinforcement during the experiment expected to get candy at the conclusion of the session.

There is a temptation to look at the many non-significant trends in the data and to cull from them a more complete discussion than is here included. However, it may be more productive to look at these trends in terms of leads for future research than in terms of their significance in this study. Social reinforcement and its components remain to be further explored since no particular conclusions concerning vocal and smiling reinforcers alone can be drawn. One qualification included in this study which needs to be varied is the sex of the experimenter delivering social reinforcement. Sex of subject was not counter-balanced, and the number of females in each group was too small to attempt an analysis of sex differences. However, there are enough data from other sources regarding the influence of the sex of the experimenter (*4*, *16*) to warrant a cautious look at the present results. That is, E_1 as a male may have had differential reinforcing properties for each of the sexes; E_1 may have had peculiar reinforcing properties unique to him alone. Thus, further investigation counterbalancing for sex of subject and sex of experimenter must follow. In addition, while the behavior of the experimenter delivering reinforcement was relatively standardized, replication over experimenters must be explored.

Percentage of reinforcement. One question which this study set out to investigate was whether or not a partial reinforcement effect could be demonstrated with mentally retarded children. This refers to an eventual same or higher level of acquisition performance and greater resistance to extinction for partially reinforced subjects when compared to continuously reinforced subjects. It cannot be said that the partial reinforcement effect is strongly evident in the data. While certain trends were present which suggest a partial reinforcement effect, most were non-significant. However, there were some clear differences between the two groups.

In no case did 50% subjects reach criterion in acquisition. It might be concluded that under 50% partial reinforcement learning did not occur. Fifty per cent subjects made more correct responses than 100% subjects during acquisition but did not reach a criterion of five consecutive correct responses. The 100% subjects, however, received significantly more reinforcements. But the groups did not differ during extinction. Thus, performance during extinction was neither a function of amount of reward nor number of correct responses during acquisition.

Both the circumstances of this experiment and the general effects of partial

reinforcement may account for the performance of 50% subjects during acquisition. It will be recalled that pilot study evidence suggested that some continuous reinforcement for any response was necessary in order to get subjects to respond on every trial. E_2's instructions were: 'Just one is the right answer.' But for the first five trials all of the three possible responses were reinforced. Thus, a contradiction between instructions and procedure existed initially for all subjects. For 100% subjects, however, after the first five trials only the correct response was rewarded. But it was rewarded on every occurrence, and in the majority of cases the subjects learned to criterion. In the case of 50% subjects only every other correct response was rewarded after the first five trials, and in no instance did a subject reach criterion. This striking difference between 50% and 100% subjects is one of the most interesting findings of the study. Fifty per cent subjects typically began by giving responses in serial order: for example, 'bird, cat, dog'; and many continued to repeat this pattern. It was noticed that several of the 50% subjects repeated this patterning throughout acquisition and into extinction.

This trend to patterning stimulates some theoretical speculation. Amsel (*1*) and Spence (*11*) have theorized that partial reinforcement creates more frustration for the learner than does continuous reinforcement. This reasoning relates to the hypothesis that nonreward produces frustration. Frustration is considered to be a drive or motivational variable, such that its presence increases the drive level of the learner. Thus, subjects under partial reinforcement would be under a higher drive level, or be more highly motivated. It is also hypothesized that the higher the drive, the more likely it is to generalize over a wider range of responses. A reinforcer would also generalize and affect a wider range of responses. If partial reinforcement conditions thus increased drive level and if patterning occurred, it may be that the reinforcer was not only reinforcing the response which immediately preceded the reinforcer, but such a chain of responses immediately preceding the reinforcer. In this case the chain of responses immediately preceding the reinforcer was the pattern of serial responses, dog, cat, bird. Such a generalized reinforcing effect would explain the persistence of the patterning throughout acquisition. Trials did proceed rapidly one upon another so that the chain of responses in patterning was relatively close together in time. If such findings can be replicated, a further problem to be explored is whether similar results are to be found with non-retarded populations as well.

Aside from theoretical aspects, the main conclusion of this study is that a partial reinforcement effect was not significantly in evidence here. Questions which remain to be answered involve the effect of a constant number of trials and/or reinforcements, instructions consistent with procedures, and a longer period of continuous reinforcement before switching to partial reinforcement. Other percentages of reinforcement also remain to be studied to determine the relative effectiveness of partial reinforcement at different percentages.

This study may hint at some possible practical implications for teaching mental retardates. If the results stand up under further research efforts, then it would seem that partial reinforcement at the 50% level does not lead to more efficient learning over a relatively short

period of time. Though partial reinforcement has been found to be effective in other situations, the learning of a complex vocal task for mental retardates may require longer periods of continuous reinforcement before switching to partial reinforcement or higher percentages of partial reinforcement. Such a generalization from the data of this study to the more complex teacher-student, clinician-patient situations in which the retardate is to be found, is made cautiously. Its validity remains to be demonstrated by future research evidence.

Summary

The present study was designed to answer questions concerning the types of reinforcing stimuli that were most effective in increasing the frequency of a correct vocal response among retarded subjects, the nature of the effectiveness of vocal and non-vocal components of social reinforcement, and the demonstrability of the partial reinforcement effect with retarded subjects.

Five reinforcing stimuli—candy, vocal, smiling, candy-vocal, and vocal smiling—were studied under conditions of partial (50%) and continuous (100%) reinforcement. Sixty mentally retarded children ranging in IQ from 36 to 78 and having PLS scores from 53.6 to 70.0 were randomly assigned to treatment and reinforcement groups.

The combination of candy-vocal reinforcers tended to be most effective. No particular conclusions were drawn concerning the relative effectiveness of vocal and smiling separately or in combination.

A partial reinforcement effect was not clearly demonstrated by the study. Certain important procedural aspects and trends in the data, however, point to interesting research problems concerning partial reinforcement. The questions which the study was designed to answer thus remain to be further explored.

References

1. Amsel, A., The role of frustrative nonreward in non-continuous reward situations. *Psychol. Bull.*, 55, 1958, 102-119.
2. Barnett, C., Pryer, M. W., and Ellis, N. R., Experimental manipulation of verbal behavior in defectives. *Psychol. Reports*, 5, 1959, 593-596.
3. Grosslight, J. H., and Child, I. L., Persistence as a function of previous experience of failure followed by success. *Amer. J. Psychol.*, 60, 1947, 378-387.
4. Hartup, W. W., Sex and social reinforcement effects with children. Paper presented at 1961 meetings of American Psychological Association.
5. Jenkins, W. O., and Stanley, J. C., Partial reinforcement: A review and critique. *Psychol. Bull.*, 47, 1950, 193-234.
6. Kanfer, F. H., Verbal conditioning: Reinforcement schedules and experimenter influence. *Psychol. Reports*, 4, 1958, 443-452.
7. Krasner, L., Studies of the conditioning of verbal behavior. *Psychol. Bull.*, 55, 1958, 148-161.
8. Lewis, D. J., Partial reinforcement: A selective review of the literature since 1950. *Psychol. Bull.*, 57, 1960, 1-28.
9. Lindquist, E. F., *Design and Analysis of Experiments in Psychology and Education.* New York: Houghton-Mifflin, 1956.
10. Rheingold, H. L., Gewirtz, J. L., and Ross, H. W., Social conditioning of vocalizations in the infant. *J. comp. physiol. Psychol.*, 52, 1959, 68-73.
11. Spence, K. W., *Behavior Theory and Learning: Selected Papers.* Englewood Cliffs, New Jersey: Prentice-Hall, 1960, Chap. 6.
12. Spiker, C. C., Performance on a difficult discrimination following pretraining with distinctive stimuli. *Child Developm.*, 30, 1959, 513-521.
13. Spradlin, J. E., Effects of reinforcement schedules on extinction in severely retarded children. *Amer. J. ment. Defic.*, 66, 1962, 634-640.

14. STEVENSON, H. W., and CRUSE, D. B., The effectiveness of social reinforcement with normal and feebleminded children. *J. Pers.*, 29, 1961, 124-135.
15. STEVENSON, H. W., and FAHEL, LEILA S., The effect of social reinforcement on the performance of institutionalized and non-institutionalized normal and feebleminded children. *J. Pers.*, 29, 1961, 136-147.
16. STEVENSON, H. W., and KNIGHTS, R. M., The effectiveness of social reinforcement after brief and extended institutionalization. *Amer. J. ment. Defic.*, 66, 1962, 589-594.
17. ZIGLER, E., The effectiveness of social reinforcers and the structure of the child. Paper presented at 1961 meetings of the American Psychological Association.
18. ZIGLER, E. F., HODGDEN, L., and STEVENSON, H. W., The effect of support on the performance of normal and feebleminded children. *J. Pers.*, 26, 1958, 106-122.

2. The Effects of Feedback Modification on Verbal Behavior

ROSS H. COPELAND

Traditional investigations of delayed auditory feedback have focused upon the vocal disturbances and speech disruption which occur in the subject's responses during the presentation of feedback delays. In typical delayed feedback studies the subject was fed back a delayed recording of the phonemes or morphemes which he was still in the process of producing. The use of 'delayed feedback' as a facilitator to speech production has not been investigated with mentally retarded or emotionally disturbed children. In two studies which employed retarded or disturbed children (*1*, *4*), the disruptive effects of delayed feedback were studied. The length of delay in traditional feedback studies has been .15 seconds. Since the delay period used in this study was greatly extended (1.0 second), the procedure was designated as *immediate feedback*. This designation had been suggested by Fairbanks (*3*).

Informal observations suggested that echoic verbal behavior of an adult tended to elicit increased vocalization from mental retardates who were severely deficient in language. Could, then, the echo of a child's own vocal productions influence the frequency of his subsequent vocal productions?

In pilot work related to the development of the apparatus (referred to below) observations suggested the hypothesis that immediate feedback would tend to increase the frequency of verbal behavior. The current experiment addresses three questions:

1. What is the effect of immediate feedback on the frequency of vocalizations made by retardates in a non-social situation?

2. Are there differences in the frequency of vocalizations emitted by retardates of different verbal levels in a non-social situation?

3. Does immediate feedback affect children of different verbal levels differently with regard to frequency of vocalizations?

Method

Apparatus. For the purposes of studying the facilitative aspects of immediate feedback on verbal behavior an apparatus had to be designed which would allow the child, after a specified time delay, to hear via free-field what he had just said. Such a system necessitated live microphones and loudspeakers within the same system and in the same room. A technical description of the apparatus is available (*2*). The apparatus can be

Ross H. Copeland (M.S., Purdue University, 1955) is Research Associate in the Bureau of Child Research, University of Kansas, and Director, Speech and Hearing Clinic, Parsons State Hospital and Training Center.

TABLE 3.8. Sex, age, language, and intelligence characteristics of the subjects.

Groups	*N* *Male*	*Female*	*Total*	*Mean Age Years-Months*	*PLS (vocal T score) Range*	*WISC* (FS) Mean*
High	28	16	44	14-6	58.3-70.0	60.0
Low	33	11	44	12-6	45.0-51.6	47.0
Total	61	27	88	13-5	45.0-70.0	53.5

*WISC scores available for only 49% of subjects.

set for time delays varying from 0.1 second to five seconds. The intensity of the feedback can also be controlled. The only portions of the apparatus visible to children were the loudspeakers and microphones in the sound-proof experimental room. All the recording, playback, and control devices were located in an adjoining control-observation booth.

In addition to delivering feedback, the apparatus recorded the frequency of the child's verbalizations which, in turn, triggered a counter. The counter registered only the number of one-second intervals in which the child made some vocal response. Verbalization which was fed back into the room did not activate the counter. In this procedure the same count was given a monosyllabic utterance as one full second of connected discourse. The counter recorded vocalizations during control and experimental conditions. The duration of the recording time and the duration of the free-field playback monitoring were inflexible. If the subject produced consecutive phonations in excess of one second, the delayed feedback signal via the loudspeakers presented only the first second of the production. If the subject offered further speech production during the delayed feedback of his previous verbalization, the speech production was not recorded until the end of the one-second delayed feedback.

Subjects. A total of 88 children participated in the experiment. Table 3.8 indicates the number, mean age, PLS range, and mean Wechsler Intelligence Scale for Children (WISC) scores of the children. The total age range of the children was seven years, five months to 16 years, seven months, with a range in the high group of 11 years, two months to 16 years, seven months and in the low group of seven years, five months to 16 years, seven months. PLS vocal T scores (described in Section One) were used as the selection criteria for assignment to high and low level groups. Forty-four subjects were randomly selected from the high-scoring 37% and 44 subjects were randomly selected from the low-scoring 27% of children for whom PLS scores were available. Excluded subjects were replaced by other subjects randomly selected from the appropriate PLS level. Two children were excluded from the study (both in an experimental session) when they became frightened (cried and pounded on the glass between the control room and the experimental room) when left alone. It should be noted that in the final group of subjects, WISC scores were not available (usually

because at time of testing the child was not considered testable) for 11.3% of the high level subjects and 90.9% of the low level subjects or for a total of 51.1% of all subjects.

Procedure. One half of the subjects from the high and low level groups were randomly designated as receiving control conditions first, and the remaining half received experimental conditions first. Subjects were randomly chosen for testing from the total group so that not all subjects receiving a given order of treatment would be run consecutively. Control and experimental conditions differed only in that no electronically mediated feedback was given during the control session.

For the experimental session, each child was escorted into a sound-proof room by the experimenter (E)[1] for an interview period which never exceeded three minutes. During this time E established rapport by talking about some toy objects in the room: E said, 'Today I want to see just how loudly you can talk. In a big, loud voice tell me your name.' When a subject had responded, E continued: 'That was just fine. Now in a big, loud voice tell me what cottage you live in.' After the subject responded, E said: 'That was fine, too. Now I have some toys I would like for you to look at and to tell me what they are. So in a big, loud voice, you tell me what the names of these things are.' As E began to present the stimuli, prior to subject's first response, the feedback was switched on so that the subject was able to hear free-field both himself and E under the delay condition at an intensity level of approximately 70 db. When the interview was completed, E announced that he would have to leave the room for a few minutes but that he would return shortly and that he wanted the subject to wait for him. Having obtained the subject's agreement to wait for him and repeated the fact that he would be back, E left the room and closed the door. E remained out of the room for five minutes. During this time free-field feedback was operative. At the end of the period, feedback was deactivated, E reentered the room and concluded the session.

In the control condition the procedure was exactly the same except that no feedback was delivered either during the interview portion or the five-minute solitary period.

A second session was administered four days following the child's first session. If the child had had the experimental condition his first session, the control condition was then given and vice versa.

Results

The total number of verbal productions during the five-minute period was available for each subject from the counter record.[2] Table 3.9 indicates the mean number of verbal productions by treatment group and order at each level of PLS score. Order 1 involved experimental condition first and control second; order 2 vice versa.

[1] In all cases E was a male college student.

[2] The sensitivity of the instrument allowed non-vocal as well as vocal noises to be registered. Two trained listeners counted non-vocal noises from the tapes and subtracted this number from the subject's total count. The judges demonstrated .98 reliability in their count. Where disagreement did occur, split differences were recorded. In this manner only vocal productions of the subjects were included in the analysis.

TABLE 3.9. Mean number of vocalizations of high and low level children under experimental and control conditions.

Level-Order Groups	*Experimental*	*Control*
High 1st Order	14.13	7.32
High 2nd Order	9.86	5.45
Low 1st Order	32.00	17.88
Low 2nd Order	32.72	15.31

A Lindquist Type III analysis of variance (*5*) was used involving two between factors (order and PLS level) and one within factor (treatment). The analysis (Table 3.10) revealed the effect of immediate feedback to be significant at beyond the .05 level. Also there was a difference between the number of vocalizations of the high and low level groups under both conditions significant at beyond the .01 level. There were no significant effects of order, order × PLS level, order × treatment, treatment × PLS level, or order × treatment × PLS level. Figure 3.5 shows the mean verbal productions by level and treatment.

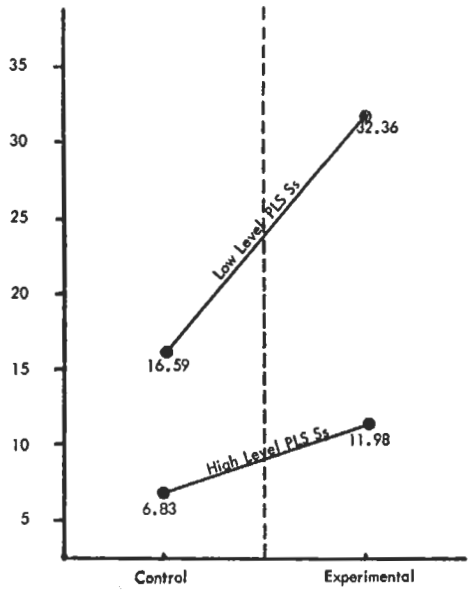


FIGURE 3.5. Mean number of vocalizations emitted by high and low level children under control and experimental conditions.

Discussion

These two findings prompt some questions and speculations. There are several possible factors involved in the re-

TABLE 3.10. Analysis of variance for number of vocalizations.

Source	*df*	*ss*	*ms*	*F*	*P*
Between Ss	87	74,522.73	.....	.....	.....
Between Levels	1	10,355.11	10,355.11	13.57	<.01
Between Order	1	2.27	2.27	.....	
Level × Order	1	90.21	90.21	.....	
Error Between Ss	84	64,075.14	762.79	.....	
Within Ss	88	68,080.00			
Between Conditions	1	5,020.45	5,020.45	6.83	<.05
Conditions × Levels	1	1,150.57	1,150.57	1.56	ns
Conditions × Order	1	176.01	176.01	.....	
Conditions × Level × Order	1	54.56	54.56	.....	
Error Within Ss	84	61,678.41	734.26		
Total	175	142,602.73			

.05 = 3.96
.01 = 6.96

inforcing effect of immediate feedback. First, there is the possibility that the increase in frequency of responses is due to the novelty of the stimulation. Perhaps any novel stimulus which was made contingent on vocalization would serve to increase the frequency of the response during a brief five-minute period. Second, the consistency of feedback may be the important factor in the increase. That is, consistent control of an aspect of the environment may in itself serve as a reinforcer. Third, the aspect of the feedback situation which increases the vocal response may be that each response yields a different but reinforcing consequence. That is, the child's response 'eee—' consistently yields feedback 'eee—' which is different from what the consistent feedback of the child's response of 'ooo' would yield. Moore and Anderson (*6*) consider this an important factor in the automated teaching process they used to teach very young children (three to five years of age) to read and typewrite.

The finding that under both control and experimental conditions low level children emitted more vocalizations than did high level children may appear puzzling at first glance. However, one should remember that a high PLS vocal score is obtained by giving an intelligible, appropriate verbal response in a social context while a high vocal score in the current experimental and control situation was based only on frequency of vocal noises. Moreover, the PLS test situation is a social situation while the current experiment counted vocalizations in a non-social setting. A possible explanation is suggested by the discussion above. Vocal behavior of high level children may be primarily under the control of social variables while the vocalization of low level subjects may not be.

The measure used in this experiment yielded significant differences between the experimental and control treatment and between the 'high' and 'low' children. However, more sensitive measures might allow more detailed analyses of the vocal behavior occurring in the feedback situations. Some areas and questions which could be investigated are the following:

1. What are the contextual aspects of the subjects' verbal productions in response to immediate feedback? It was informally observed that many of the children offered quite hostile verbalizations while 'playing with their own echoes.'
2. What is the rate of extinction for subjects who have demonstrated high levels of performance under feedback modification?
3. What are the differential effects of immediate feedback incorporating delays of two and three seconds instead of one second?
4. What are the differential voice variables of the responses made under conditions of immediate feedback, particularly those of rate, phonation/time ratio, speech intensity, and pitch?

The most important finding of this study is the demonstration of the influence of electronically mediated, echoic variables on verbal behavior. Subsequent research remains to be done in order to understand the process involved in this phenomenon as well as to describe the relevant variables more definitively.

Summary

Free-field delayed feedback was administered to 44 high and 44 low verbal level subjects in a solitary situation in

order to determine the effect on verbal behavior It was found that the feedback condition elicited a significantly greater amount of verbalization from both the high and low level groups. It was additionally noted that the low level subjects verbalized with significantly greater frequency under both experimental and control conditions than did the high level subjects.

References

1. Barnes, N. S., Delayed feedback: voice variable alterations related to intelligence levels of mentally retarded adults. Unpublished thesis, Purdue University, 1957.
2. Copeland, R. H., and Halpain, N., A device for free-field monitoring of delayed speech feedback. *J. Speech Hearing Dis.*, 26, 1961, 137-140.
3. Fairbanks, G., A device for continuously variable time delay of headset monitoring during magnetic recording of speech. *J. Speech Hearing Dis.*, 16, 1951, 162-164.
4. Goldfarb, W., and Braunstein, Patricia, Reactions to delayed auditory feedback in schizophrenic children. In P. H. Hock and J. Zubin (editors), *Psychopathology of Communication.* New York: Grune and Stratton, 1958, 49-63.
5. Lindquist, E. F., *Design and Analysis of Experiments in Psychology and Education.* New York: Houghton-Mifflin, 1956.
6. Moore, O. K., and Anderson, A. R., *Early Reading and Writing, Part I: Skills.* (Motion picture) Basic Education, Inc., Guilford, Connecticut, 1960.

Summary and Overview

JOSEPH E. SPRADLIN

Rationale

The preceding three sections have presented research on test development, two-person language interaction, and the effects of consequences on vocal behavior of retarded children. As stated in the introduction, three assumptions have guided the research: first, the study of language and communication is facilitated if terms are defined so that the events to which they refer can be observed, measured, and classified; second, language and communication are determined by variables which can be objectively described, classified, and, in many instances, manipulated; finally, language is learned behavior and as such is subject to the same principles that apply to other forms of learned behavior.

The impact of these assumptions on the current project is to be found in the kinds of research problems posed and, equally significant, in the manner in which this research is discussed. Consistent with the first two assumptions, a studious attempt has been made in this monograph to avoid characterizing language and communication in terms of inferred or implicit processes. Rather, *language* has been defined in terms of the observable responses of people, with reference to such specific behaviors as vocalizations, gestures, or writing.

The conception of language as a form of learned behavior, the third of these assumptions, permeates the research reports, even where 'formal' learning studies have not been conducted. The PLS emanates directly from a learning frame of reference, and it is assumed that the responses sampled by this instrument are subject to learning principles, that is, they can be modified through reinforcement and discriminative stimulus training. Behavior occurring in interpersonal settings is also considered learned, and subsequent research may investigate the function which such reinforcements as social approval, vocal reply, reprimand, or object reinforcers play in maintaining and directing communication between persons. The studies of consequences reported in the final section of this monograph reflect even more directly the application of learning principles to the language behavior of mentally retarded children. The extension of these studies may help us better understand the relative effectiveness of various classes of reinforcers on the vocal behavior of these children and, further, may indicate the extent to which knowledge about schedules of reinforcement, extinction, response shaping, and discrimination applies to the language behavior of retarded children.

Conclusions and Indications for Additional Research

One of the salient features of the Parsons Language Sample (PLS) described in Section One is that it is an objective instrument that does not require an examiner with extensive training or background for its administration. Freshman college students, after approximately a week of training, obtained equivalent score distributions when testing random samples of the same population of children. Further, when they tested the same children, inter-examiner reliability was quite high. The scores on the PLS (excluding those for the mand subtest) were highly stable for periods of up to six months except for a group of newly institutionalized boys who were tested shortly after initial institutionalization and retested 15 months later. Instability of test scores for these new entrants may have been a function of the extended test-retest interval, or it may be a function of changes in performance of the children that accrue to the 'institutionalization process.'

Subtest comparisons revealed that the three vocal subtests and the comprehension and echoic gesture non-vocal subtests were highly intercorrelated. It may not be useful, therefore, to treat these subtests independently in subsequent applications of the PLS. On the other hand, the intraverbal gesture subtest (gestural responses to oral questions) was uncorrelated with the other five subtests and apparently samples behavior that is distinct from that measured by the remaining portions of the test. The relatively gross dichotomizing of children as 'low' or 'high' on the basis of the composite vocal section of the PLS proved useful in a number of experimental investigations. Children dichotomized on this basis consistently differed in verbal behavior in small group settings. Moreover, both the vocal and non-vocal subsections of the PLS correlated significantly with ranks assigned the children by cottage aides for speech and non-speech communication.

In summary, the PLS gives evidence of being a highly reliable instrument which is relatively simple to administer and which has validity for predicting non-test language behavior of retarded children. In future work the range of behavior sampled by the PLS should be extended to include more items that discriminate among severely retarded children. Though the test in its present form is considerably more sensitive to low levels of verbal behavior than is the WISC or the Binet, 87 of 275 children tested failed to make any appropriate response to the vocal section and 22 subjects made no appropriate responses to any items within the test. Item difficulty should be carefully studied since this will not only allow for evaluation of the test but may also give some indication of the sequence in which various behaviors are learned. Another attempt should be made to devise items that sample the kinds of social demands and requests included in the unreliable mand subtest. Procedures for scoring the entire test should be refined, particularly to take advantage of the negligible relationship between the intraverbal gesture subtest and the remaining portions of the PLS. An attempt should be made to determine what non-test behaviors relate to the intraverbal gesture subtest. For example, do children who score high on this subtest use more gestures in ordinary social situations than do low scorers? What communication function is

served by gesture? What variables control gestures? Finally, it is important to learn whether particular patterns or 'performance profiles' on the PLS relate in any predictive way to non-test behavior and, more significantly, to verbal learning among these children.

The interpersonal studies reported in Section Two provided information concerning behavior of retarded children and of adults with whom they were assembled. High level children consistently emitted greater and more complex vocal output than did low level children when the children were dichotomized initially on PLS performance. Both groups of children used fewer words in a 'tutorial' situation with adults than they did in less structured, informal situations with the adults. Children exposed to an 'interview' treatment did not differ significantly in verbal behavior from children who were given a supportive, 'clinical' treatment.

There was some evidence that adult verbal behavior was affected by (a) verbal level of the children with whom they were assembled, (b) the type of situation in which they were placed, and (c) the kinds of preinstructions they were given. When teaching a task, adults spoke more often, used more words, had greater mean length of response, asked fewer questions, and used smaller type-token ratios than when they were instructed merely to take care of the child. Adults who 'interviewed' children resorted largely to questions. They did not differ significantly from adults in a 'clinical' treatment with regard to number of responses, number of words, or MLR. Type-token ratio consistently differentiated adult assemblies with low versus high level children. Regardless of the type of experimental situations, adults invariably used smaller TTR's with low level children. Adults also differed in response to the two levels of children on measures of the number of questions, but not consistently from study to study.

In summary, the interpersonal studies suggest that the verbal behavior of adults was at least partially under control of the verbal characteristics of the retarded children with whom the adults were assembled. In addition, adult behavior varied in relation to the experimental situation as well as to pre-experimental instructions. Although the interpersonal paradigm developed in this monograph dealt only with adult-child assemblies, the paradigm is equally appropriate for the study of child-child (*3*) or adult-adult (*1*, *2*) groupings. The participants may be assembled on the basis of diverse 'naturalistic' or experimentally determined variables. Once in the assembly, behavior of either or both participants may be allowed to vary freely, or attempts may be made to direct behavior by training or instructions. Sessions may be set up on competitive or cooperative bases, and the experimental task may be highly structured or very informal and may vary considerably in complexity.

Horowitz' study, reported in Section Three, compared the effects of type and percentage of reinforcement on the acquisition and extinction of a complex verbal task. Five types of reinforcement (candy, social, vocal, vocal plus social, candy plus vocal) and two schedules of reinforcement (50% and 100%) were studied. The types of reinforcement were neither significantly different in the number of trials necessary to make five consecutive correct responses nor in response latency during acquisition. Children receiving social reinforcement,

however, reached criterion in the least number of trials during extinction while the candy and the vocal subgroups required the greatest number of trials before extinction was reached. None of the subjects in the 50% reinforcement conditions reached the criterion of five correct consecutive responses during the acquisition (or 'learning') phases of the experiment. In contrast, 57% of the subjects in the 100% reinforcement condition successfully learned the task. Although subjects in the 50% group did not learn to criterion during acquisition training, the percentage of correct responses for this group, as well as for subjects in the 100% reinforcement group, increased during acquisition training. There was some evidence to suggest that subjects in the partial reinforcement group were more resistant to extinction than those in the continuous reinforcement group, but this result was not conclusive.

The comparative effectiveness of various types of schedules of reinforcement on the learning of verbal tasks by retarded children should be explored much more extensively. The relevance of principles of learning derived from studies of normal children to learning of retardates warrants careful and systematic exploration. To the degree that principles of generalization, discrimination, reinforcement, etc., apply to retarded children, a body of knowledge will be available as a source of basic understanding. The Horowitz study should lead to additional investigation of the effect of partial reinforcement on complex verbal learning by mental retardates. The 'patterning' of responses demonstrated by the partially reinforced subjects is particularly intriguing and deserving of further study.

The second study reported in Section Three dealt with the effect of freefield delayed feedback on the vocal behavior of children with high and low language skills. During a feedback condition, vocal responses of the subject which occurred during alternate one-second periods were 'echoed' by a freefield sound system one second after the vocalization occurred. This treatment resulted in an increase in the number of vocal responses for children of high and low verbal levels. Low level children in experimental and control conditions issued more vocalizations than did the high level subjects. This finding may suggest that the vocal responses of higher and lower level children may differ as a function of presence or absence of a listener. It is quite possible, of course, that any mild and novel auditory (or other) stimulus would have served as well to increase vocalization of the children in the experimental condition. Further, it is not known to what extent the temporal arrangement between the feedback and the child's response is critical. That is, the feedback may have served either as a source of reinforcement (feedback after the response) or as the occasion or 'discriminative stimulus' for the child's vocalization (feedback before the response).

Although the research reported in this monograph has been presented in discrete sections, in many ways the areas converge and will continue to do so as more research within the Parsons Project is completed. The development of the PLS reported in Section One was vital for the selection and classification of subjects used in the studies reported in Sections Two and Three. In turn, the concepts of interpersonal effects discussed in Section Two are clearly ap-

plicable to the kinds of examiner-subject interactions that characterize the testing process used with the PLS. Similarly, while the PLS derives from a learning theory framework, at least one of the studies reported in Section Three is a specific test of the applicability of a similar framework to retarded children. In future studies, the areas included in the current monograph, and additional ones, should be even more closely related. Certainly it would seem appropriate to investigate whether particular patterns of response on the PLS have any relation to subsequent behavior in interpersonal or in learning situations. It would also seem worthwhile to consider the effects of specific reinforcement parameters, such as those included in Section Three, in a variety of interpersonal situations. In this context it would be appropriate to investigate the reinforcing properties of verbal imitation (echoing) when the 'echo' is provided by another person rather than by the auditory system used by Copeland. In more general terms, knowledge of procedures for assessing and modifying verbal behavior of retarded children in 'communication' circumstances will undoubtedly involve the integration of several sources of experimental inquiry, including, it is hoped, those considered in this monograph.

References

1. Jaffee, J., Language of the dyad: a method of interaction of analysis in psychiatric interviews. *Psychiatry*, 21, 1958, 249-258.
2. Matarazzo, J., Saslow, G., and Matarazzo, Ruth, The interaction chronograph as an instrument for objective measurement of interaction patterns during interviews. *J. Psychol.*, 41, 1956, 347-367.
3. Rosenberg, S., Spradlin, J. E., and Mabel, S., Interaction among retarded children as a function of their relative language skills. *J. abnorm. soc. Psychol.*, 63, 1961, 402-410.

Appendix A

Test Booklet, Parsons Language Sample

Name_____ Birthdate_____ Date_____

Key

√ = appropriate
× = inappropriate
⊗ = unintelligible
— = no response

1. *Tact:* The examiner shows the child each object or picture one at a time. Upon presenting each item the examiner will say, 'What is it?' or 'What do you call it?' If the child does not respond the examiner repeats the question. The examiner accepts and rates any response given by the child. After the child has responded, the examiner says, 'Good' or indicates approval to the child, then hands the child the toy for inspection. As soon as possible the examiner retrieves the object and proceeds to the next item. Responses are rated according to the above key.

		Vocal Rating	Correct responses
(Real Objects)	1. Ball		Ball
	2. Cup		Cup
	3. Telephone		Phone or telephone
	4. Spoon		Spoon
	5. Pencil		Pencil
	6. Wrench		Wrench
	7. C Clamp		Clamp or C clamp
(Miniature Objects)	8. Duck		Duck
	9. Car		Auto, car, or automobile
	10. Chair		Chair
	11. Table		Table
	12. Screwdriver		Screwdriver
	13. Pliers		Pliers
	14. Bottle Brush		Brush or bottle brush
(Colored Pictures)	15. Motherly type woman		Mother, lady, or woman
	16. Kitten		Cat, kitty, kitten, or kitty cat
	17. Apple		Apple
	18. Drum		Drum
	19. Nail		Nail
	20. Leaf		Leaf
	21. Anchor		Anchor

(Tact Continued)

		Vocal Rating	Correct responses	
(Non-colored Pictures)	22. Fatherly type man		Man, daddy, father, or workman	
	23. Puppy		Dog, doggie, or puppy	
	24. Baseball bat		Bat or baseball bat	
	25. Watch		Watch	
	26. Arrow		Arrow	
	27. Feather		Feather	
	28. Propeller		Propeller	
	Mand item #1 a. The examiner holds the wind-up toy in the child's view. b. The examiner winds up the toy and places it on the table out of the child's reach and allows it to run five seconds. c. The examiner picks up the toy and again holds it in the child's view for ten seconds. If the child has not responded within this time, the object is put back in the drawer. If the child asks for the toy, a mand is recorded. Gestures are also recorded.	Gesture	Vocal	Correct response Vocal – Any vocal request for the toy. Gesture – Any gestural request for the toy. Merely reaching for the toy is not classed as a gesture.

2. *Echoic:* Prior to administering each item the examiner makes certain he has the child's attention. This is necessary since the examiner says the word and numbers only ONCE. If any item is passed, the following two items must be administered in each section. After three consecutive failures in first section of the echoic test, discontinue that section and proceed to the next section. Correct response consists of repetition of the stimulus presented.

	Vocal Rating
Section A 1. Say 'ball.'	
2. Say 'skate.'	
3. Say 'cowboy.'	
4. Say 'playhouse.'	
5. Say 'Give me one.'	
6. Say 'The cat is black.'	
7. Say 'Bob made a box for his cat.'	
8. Say 'My sister wants Daddy to buy her a big doll.'	
9. Say 'At night we went to see a movie at the theater.'	
10. Say 'In the summertime the little children like to eat black walnut ice cream.'	
Section B. 1. Say '2'	
2. Say '1'	
3. Say '3-9'	
4. Say '1-4'	
5. Say '4-9-3'	
6. Say '4-6-3'	
7. Say '9-7-6-8'	
8. Say '2-8-1-3'	
9. Say '5-4-8-7-1'	
10. Say '3-9-6-7-1'	
11. Say '1-7-9-3-2-5'	
12. Say '1-5-8-9-3-7'	
Mand item #2 The examiner pounds a peg in the pegboard, then he hands the board to the child and says, 'You do it.' The examiner retains the mallet. A mand response is recorded if the child requests the mallet vocally.	

3. *Echoic Gesture:* The gesture will be demonstrated by the examiner three times for all items. If the child is successful on echoing the gesture on any of the three trials, he is given credit. Correct response consists of imitation of the examiner's motor response.

	Gesture
1. The examiner points toward the light and says, 'Do this.'	
2. The examiner doubles up fist and pounds desk twice saying, 'Do this.'	
3. The examiner claps hands and says to the child, 'Do this.'	
4. The examiner shakes his head and says to the child, 'Do this.'	
5. The examiner rubs top of head with palm of hand and says, 'Do this.'	
6. The examiner slaps left knee with left hand and says, 'Do this.'	
7. The examiner slaps left knee with right hand and says, 'Do this.'	
8. The examiner places one Kohs block on the table, taps it with his finger and says, 'Do this.'	
For items 9-13, the examiner places two Kohs blocks on the table and taps them with his fingers according to the right (R) left (L) sequences listed. Correct response consists of imitation in exact sequence.	
9. The examiner taps on blocks R-L-L.	
10. The examiner taps on blocks R-L-R-R.	
11. The examiner taps on blocks R-R-L-R-R.	
12. The examiner taps on blocks L-R-R-R-L-R.	
13. The examiner taps on blocks L-L-R-R-L-L-L.	

Mand item #3		Vocal	Correct Response
The examiner makes the battery operated car go and then hands it to the child saying, 'You make it go.' The examiner keeps the controls. The child should request the control in some way for mand to be scored.			Vocal – A vocal request for the controls. Gesture – A nonvocal request for the controls. Merely reaching is not classed as a gesture.

4. *Comprehension:* The purpose of this subtest is to sample the child's comprehension of three types of directions—Vocal directions only, gestural directions only, and vocal and gestural directions given together. With the exception of the first item it is necessary to obtain the child's attention before administering the item. All items may be repeated *once* and *once only* if the child does not respond to the item the first time the directions are given. The item must be repeated exactly as it was initially given. Correct response consists of execution of the command.

Directions	Gesture
1. (Vocal direction only.) The examiner waits until the child is looking away from him and then says the child's name. If the child changes position so that he is looking toward the examiner, credit is given.	
2 (Vocal AND gestural directions.) The examiner gets up from his chair and moves away from the child. When he is standing from 6 to 10 feet away from the child he extends his arm and makes a beckoning motion with his hand and arm and says, 'Come here.' Credit will be given if the child moves toward the examiner.	
3. (Vocal only.) The examiner says, 'Open the door.' Once the door is open the examiner says, 'Now close the door.' Either part of the directions may be given twice. If the child does not respond correctly to part 1, part 2 is not given.	
4. (Gesture.) The examiner points to the child and then to the child's chair saying nothing. If the child responds by taking his chair, credit will be given. If after two administrations of the command the child does not return to his seat, the examiner vocally tells him to do so but credit is not given.	
5. (Gesture only.) The examiner places a ball, doll, and a mechanical toy (examiner's left to right) on the desk close to the child. If the child picks up any of the toys, the examiner holds out his palm just in front of the toy with his palm	

(Comprehension continued)

Directions	Gesture
up. Credit is given if the child places the toy in the examiner's hand. After finishing the item, the examiner removes objects from the desk.	
6. (Vocal only.) The examiner says, 'Put your finger on your nose.' Credit is given if child places finger on his nose.	
7. (Vocal only.) The examiner places a cup and spoon on the desk and says, 'Put your finger on the cup.' Credit is given if the child puts his finger on the cup.	
8. (Vocal only.) The examiner says, 'Put the spoon in the cup.' Credit is given if the child places the spoon in the cup.	
The examiner places the following objects in front of the child (examiner's left to right): Cup, spoon, toy car, toy purse, toy chair. After each item the objects are returned to the above positions. The items are left on the table for items 9 through 18.	
9. (Vocal AND gesture.) The examiner says, 'Put the cup to your mouth' and makes a gesture as if he were picking up a cup to put to his (the examiner's) mouth.	
10. (Gesture only.) The examiner points to the chair and then to the top of the purse. Credit is given the child if he places the chair on or in the purse.	
11. (Vocal only.) The examiner says, 'Put the car on the floor.'	
12. (Vocal AND gesture.) The examiner says, 'Put the cup beside the chair' while pointing first to the cup then to a spot to the right of the chair. Credit is given if the cup is placed in the commanded spot.	

(Comprehension continued)

Directions	Gesture	Correct Response	
13. (Gesture only.) The examiner points first to the cup then to the car and holds his upward palm. Credit is given if the child places both objects in the examiner's hand.			
14. (Vocal only.) The examiner says, 'Put the spoon in the purse; put the car on the floor.' Credit is given if task is completed according to the above sequence.			
15. (Vocal AND gesture.) The examiner says, 'Put the chair in the cup' (while pointing to the chair then the cup) then, 'Put the spoon beside the purse' (while pointing first at the spoon then at a spot to the right of the purse). Credit is given if command is followed in sequence.			
Mand item #4 The examiner reaches in the drawer (obtains three buttons) then puts his closed hand in front of the child and says, 'Guess how many buttons I have in my hand?' If the subject guesses, examiner will put the buttons back in the drawer and wait for 30 seconds. A response is recorded if the subject indicates in some way that he wants feedback on his rightness or wrongness. The guess is not scored; however, the demand for feedback is.		Vocal	Vocal—A vocal request to know how many buttons were in the examiner's hand. Gesture—Any gestural request for feedback.
16. (Vocal only.) The examiner says, 'Give me the car, place the spoon under the chair, then put the purse on the floor.' Directions must be followed in sequence for credit.		Execution of the command.	
17. (Vocal only.) The examiner says, 'Put the car in the purse, put the spoon on the floor, and put the cup to your mouth.' Directions must be followed in sequence for credit.		Execution of the command.	
18. (Vocal only.) The examiner says, 'Look at the light, open the door, and put the cup to the right of the purse.' Directions must be followed in sequence for credit.		Execution of the command.	

5. *Intraverbal:* Prior to asking the child the question, the examiner makes sure he has the child's attention. Each item may be read twice. Read each item exactly as it is written.

	Vocal	Correct Response
1. What do we do when we are hungry?		Supper, get meal, tell aide, go to the dining room, or dinner.
2. Why do we have houses?		Live in, shelter, keep warm, or keep dry.
3. Why do we have stoves?		Keep warm or cook on.
4. Why do we have books?		Read.
5. Why do we have clothes?		Wear or keep warm.
6. Why do we have beds?		Sleep.
7. Before we go outside we put on our coats and ____.		Any article of clothing except coat or jacket.
8. The flag is red, white, and ____.		Blue.
9. The color of an apple is ____.		Red.
10. We go to church on ____.		Sunday or the bus.
11. Santa Claus comes on ____.		Sled, Christmas, or Christmas Eve.
12. We wear shoes and socks on our ____.		Feet.
13. We smoke ____.		Cigars, cigarettes, tobacco, or pipe.
14. Sister is a girl, brother is a ____.		Boy.
15. A lemon is sour, sugar is ____.		Sweet.
16. A car goes on the ground, an airplane goes in the ____.		Sky, air.
17. A mile is long, an inch is ____.		Small or short.
18. A chair is made of wood, a window is made of ____.		Glass.
19. Snow is ____.		White, cold, water, wet, or made of rain.
20. You kick with your foot, you throw with your ____.		Arm or hand.
21. We smile when we are happy, we cry when we are ____.		Sad or unhappy.
22. My Daddy's sister is my ____.		Aunt.
23. My Daddy's brother is my ____.		Uncle.

(Intraverbal continued)

	Vocal	Correct Response
24. In what way are a cat and a dog alike?		'Both have . . .' or 'They have . . .' any characteristic which they have in common.
25. In what way are a boat and a car alike?		'Both have . . .' or 'They have . . .' any characteristic which they have in common.
26. In what way are paper and wood alike?		'Both have . . .' or 'They have . . .' any characteristic which they have in common.
27. In what way are a tree and a lion alike?		'Both have . . .' or 'They have . . .' any characteristic which they have in common.
28. In what way are cigarettes and cigars alike?		'Both have . . .' or 'They have . . .' any characteristic which they have in common.
29. In what way is an egg and a seed alike?		'Both have . . .' or 'They have . . .' any characteristic which they have in common.

6. *Intraverbal Gestural:* The aim of this test is to measure the child's gestural behavior. Credit is given for gesture when the child either answers the question with a gesture or accompanies his vocal answer with a gesture. Prior to asking the child the question, the examiner makes sure he has the child's attention. Ask each question *only once*. Correct response consists of an appropriate gesture.

	Gesture	
1. Where is the light?		
2. Where is your ear?		
3. Can a bird fly?		
4. Can a dog fly?		
5. Can a rabbit eat?		
6. How do you fasten a button?		
7. What do you do with a cup?		
8. Can a boy outrun a horse?		
9. What do you do with a key?		
10. What do you do with a spoon?		
11. What do you do with a scissors?		
12. What do you do with a crayon?		
13. What do you do when you are hungry?		
14. What does an airplane do?		
15. What does a wheel do?		
16. What does a swing do?		
17. What do you do with a saw?		
18. What do you do with a cigarette?		
19. What is a ceiling?		
20. What do you do with a drum?		
21. What do you do with a balloon?		
22. What do you do with a comb?		
23. What is a floor?		
24. What do you do with a handkerchief?		

(Intraverbal Gestural continued)

	Gesture	Correct Response
Mand item #5 The examiner hands child a sheet of paper and says, 'Please write your name here.' The examiner makes sure that the child has no writing implement before asking the question. The item is scored for vocal if the child asks for a pencil or pen. If the child gestures for the pencil, a gesture is recorded.		Vocal—A vocal request for a pencil. Gesture—A gestural request for a pencil.

Appendix B

Standard Score Transformation Tables

Tact

Raw Score	*Standard Score*
0	30
1	31
2	32
3	34
4	35
5	36
6	38
7	39
8	40
9	41
10	43
11	44
12	45
13	47
14	48
15	49
16	50
17	52
18	53
19	54
20	56
21	57
22	58
23	60
24	61
25	62
26	63
27	65
28	66

Echoic Gesture

Raw Score	*Standard Score*
0	31
1	34
2	37
3	40
4	43
5	46
6	49
7	52
8	55
9	58
10	61
11	64
12	67
13	71

Echoic

Raw Score	*Standard Score*
0	31
1	33
2	35
3	36
4	38
5	40
6	42
7	44
8	46
9	47
10	49
11	51
12	53
13	55
14	57
15	59
16	60
17	62
18	64
19	66
20	68
21	70
22	71

Comprehension

Raw Score	*Standard Score*
1	32
2	34
3	36
4	38
5	40
6	42
7	44
8	46
9	48
10	50
11	52
12	54
13	56
14	58
15	60
16	63
17	65
18	67

Intraverbal		Intraverbal Gesture	
Raw Score	*Standard Score*	*Raw Score*	*Standard Score*
0	39	0	40
1	40	1	42
2	41	2	43
3	42	3	45
4	43	4	47
5	44	5	48
6	45	6	50
7	46	7	52
8	47	8	54
9	49	9	55
10	50	10	57
11	51	11	59
12	52	12	60
13	53	13	62
14	54	14	64
15	55	15	66
16	56	16	67
17	57	17	69
18	59	18	71
19	60	19	72
20	61	20	74
21	62		
22	63		
23	64		
24	65		
25	66		
26	68		

Appendix C

T Score Transformation Tables

Tact

Raw Score	*T-Score*
28	73
27	70
26	68
25	66
24	64
23	62
22	61
21	59
20	57
19	55
18	54
17	53
16	52
15	52
14	51
13	51
12	50
11	49
10	48
9	48
8	48
7	47
6	46
5	46
4	45
3	44
2	44
1	44
0	39

Echoic

Raw Score	*T-Score*
19	71
18	67
17	64
16	62
15	60
14	58
13	57
12	55
11	54
10	53
9	52
8	51
7	49
6	48
5	46
4	46
3	46
2	45
1	44
0	38

Echoic Gesture

Raw Score	*T-Score*
12	71
11	67
10	62
9	59
8	56
7	52
6	48
5	45
4	44
3	43
2	42
1	41
0	37

Comprehension

Raw Score	*T-Score*
18	73
17	69
16	65
15	61
14	58
13	55
12	52
11	51
10	51
9	50
8	49
7	48
6	47
5	46
4	45
3	42
2	40
1	38
0	33

Intraverbal		Intraverbal Gestures	
Raw Score	*T-Score*	*Raw Score*	*T-Score*
25	73	17	71
24	69	16	71
23	69	15	68
22	68	14	68
21	65	13	63
20	62	12	61
19	60	11	61
18	58	10	60
17	58	9	59
16	57	8	59
15	55	7	57
14	54	6	56
13	54	5	54
12	54	4	51
11	54	3	49
10	53	2	45
9	52	1	45
8	52	0	39
7	52		
6	52		
5	52		
4	52		
3	51		
2	51		
1	51		
0	43		

Appendix D

Instructions for Raters on Speech Communication

Cottage_____ Date_____ Rater_____

The Research Project is attempting to develop a language or communication test. We feel that the test has been developed to the stage now that it is fairly adequate for assessing the child's communication skill. One of the ways that we hope to evaluate our language test is by comparing the score obtained from it with the ways that persons who are well acquainted with the children rate their communication and language skills. For this reason, it is necessary to enlist your help in evaluating the children. We would like for you to rank the children on the list from best to poorest with regard to speech.

In ranking children on speech it would be well to consider the following questions:

1. When the child wants something does he ask for it, or does he gesture or pull? The child who generally asks would rank higher in speech than one who generally pulled or gestured.
2. Does the child name or describe things that happen in the cottage or on the playground?
3. Does the child answer questions with speech or does he answer with gestures or by other non-speech communication?
4. Can the child's speech generally be understood? A child who talks a lot but is unintelligible should be ranked lower than one who can be better understood.

After considering the following children with regard to the above four questions rank them from best to poorest in speech in the following manner: In the blank next to number (1) write the name of the child you feel speaks best. Next to blank (2) place the name of the child you feel speaks next best, and so on for all the listed children. Rank only the children listed on the sheet.

Appendix E

Instructions for Raters on Non-Speech Communication

Cottage_____ Date_____ Rater_____

The Research Project is attempting to develop a communication test. One of the ways that we hope to evaluate our test is by comparing the scores obtained from it with the ways that persons who are well acquainted with the children rate their communication skills. For this reason, it is necessary to enlist your help in evaluating the children. Children communicate with gestures as well as speech. We would like for you to rank the children on the list from best to poorest with regard to *non-speech* communication. *Do not make your judgment on the basis of how well the child talks.* Consider the following questions before ranking the children.

1. Does he follow directions?
2. Are you able to show the child how to do simple tasks, such as washing, dressing, handwashing, dusting, sweeping, etc.? The child who imitates your demonstrations would tend to rank higher than a child who does not.
3. Does the child communicate with you by using gestures? A child who uses many gestures would tend to be ranked higher than one who uses fewer gestures.

After considering the following children with regard to the above three questions, rank them from best to poorest in *non-speech* communication in the following manner: In the blank next to number (1) write the name of the child you feel has the highest level of non-speech communication. Next to blank (2) place the name of the child you feel communicates next best, and so on for all the listed children.

Appendix F

Interview Topics

Family

Grandparents, Aunts, Uncles, Cousins

1. Names
2. Ages
3. Occupation
4. Education Level
5. Health

Brothers and Sisters

1. Occupation (if any)
2. Education Level
3. Ages
4. Feelings of the child toward his brothers and sisters
5. Marital status of brothers and sisters (if married, how many children)

Child's Attitudes, Interests, and Feelings

1. Entertainment preferred (TV programs, movies, cartoons)
2. Favorite foods (desserts, main course, candy, drinks)
3. Favorite play (solitary play, group, baseball, etc.)
4. Feelings about being in an institution away from home
5. Feelings about the aides or cottage attendants
6. Feelings toward parents, brothers, sisters, and other significant people in the child's life
7. Feelings toward other children
8. Favorite hobbies
9. Favorite pets
10. Favorite friends in the institution
11. Favorite friends at home
12. Favorite holiday and why
13. Attitudes of child toward going home

Appendix G

Instructions to Adults

We are interested in various methods for helping children improve their communication behavior. In this particular experiment we are interested in determining whether or not having a retarded child enter into a relationship with an adult will result in improved expressive ability. The children you will see will differ in their willingness to talk and in the articulateness of their speech. All of them, however, have the need for considerable improvement in their ability to express their ideas, needs, and feelings. Some of these children do scarcely any spontaneous speaking at all and have to be encouraged to do even a minimum of talking. Others are quite shy at first and have to spend some time with an adult before they will 'open up.' Even those children who seem to have speech that is quite articulate have to be encouraged to express their ideas more accurately and completely.

In a permissive free play situation in which an adult pays considerable attention to the children, it is felt that the accuracy and amount of verbalization will tend to improve. You have complete freedom within these sessions to use whatever procedures you choose to help the children. Toys will be provided and you may use these as you wish. The only restriction is that the children must remain in the room with you during the entire session.

Each of you will be assigned four children. You will meet with each of these children once a week for 12 sessions; a session will be 15 minutes long. At the end of the session you will be given a piece of paper. On this please note your impressions of the session, techniques that seemed more or less useful, and any general impressions you may have.

Do you have any questions?

Appendix H

Prototypes of Instructions to Typists

Tape recordings have been obtained for a series of experimental sessions, each session involving an adult and a child. Each session is on a separate tape. The tapes have been randomized and divided into two groups, and it is now necessary to have typed transcripts prepared from them. Each of you will be responsible for one set of tapes. The order in which you are to type the tapes has been indicated on the last page of these instructions. Be sure to follow this order faithfully.

In preparing these transcripts or protocols, you will be asked to perform a number of functions simultaneously:

1. You will have to do a careful and accurate job of representing all the verbal activity that occurred within each session. This is extremely important since all subsequent analyses will derive from the transcripts you type.
2. You will have to differentiate the verbal activity of the child from that of the adult.
3. You will have to learn several rules concerning the designation of 'vocal response units' so that you can mark off responses on transcripts as you prepare them. You will also have to indicate whether each vocal response unit is a statement or a question.

Before discussing specific rules for marking off responses on the transcripts, I would like to present some general instructions for your consideration:

A. General Instructions:
   1. Type the transcripts in the predetermined random order.
   2. Differentiate verbalizations of the adult from those of the child by placing the identifying symbol (a) in the margin for adult verbalizations and (c) for remarks made by the child.
   3. *Do not use capitals* (except for proper names or for the pronoun 'I'), commas, question marks, or any other form of punctuation in preparing these transcripts. You will use apostrophes, however, to indicate a contraction (I'm, he's) or to indicate possession (the aide's house).
   4. Some of the remarks made by either the child or the adult will be completely or partially incomprehensible. This may be because the speaker was particularly soft-spoken, mumbled, had unintelligible speech, or because some noise obscured what the speaker was saying. If a response (to be defined later) is either *partially or completely* incomprehensible, exclude it from the transcript. Even if the response has only one incomprehensible word, leave out the entire response.
   5. Sometimes the adult or the child will make some non-communicative noises during the session. For example, the adult may say, 'The dog goes *bow-wow* and the lion goes *grr*.' If, as in the above remark, the noise is an integral part of the response, type it in. If, however, the noise is not essential, omit it. For example the child may say, '*Bow-wow*, here comes the dog.' In this instance omit the expression '*bow-wow*.'
   6. Interjections such as 'uh,' 'er,' should be omitted except when they are used as words. Examples:
      Give me the *er* book.
      Uh uh, you can't have it.
      The 'er' should be omitted.
      'Uh uh,' meaning 'no' should be typed.
   7. If the speaker starts but does not finish a word and you are quite sure what he was going to say, in-

clude the word, but place it between parentheses. For example:
I th– I know he's coming.
I (think) I know he's coming.
If you can't tell what the started word was meant to be, simply exclude it.

B. *Designating 'vocal response units.'* In this study we are concerned with the *speech* behavior of the adults and children rather than with how their responses would look on paper. We are preparing these transcripts as a convenience, but more basically we are concerned with how the individuals used speech in the actual experimental sessions. We are not interested in whether or not a given response was grammatically complete and accurate. Rather we want to know whether it was functionally complete in terms of the ongoing exchange between the adults and the children. In normal conversation we don't always have a well defined predicate and nominative; and we indicate the beginning and end of our expressions by pauses, inflections, shifts in topics, etc., rather than by commas, periods, or exclamation points. That is why we have asked you not to put these punctuation marks in the transcripts you prepare. A little later I will describe the system you will use to indicate when a *vocal response unit* begins and ends. First, let us consider some of the rules that will help you decide when such a unit has occurred.

1. In general, a vocal response unit is a unit of spoken language marked off on either side by a pause or by some change in inflection.
2. A vocal response unit is considered finished when the speaker come to a complete stop and allows his voice to fall.
3. A vocal response unit is considered finished when the speaker comes to a complete stop with either a questioning or exclamatory inflection.
4. A vocal response unit is considered finished when the speaker in some manner clearly indicates he does not intend to complete the remarks.
5. A vocal response unit is considered completed when one speaker terminates and the other begins speaking.
6. A vocal response unit may include several simple utterances. If one simple utterance or remark is immediately followed by another with no pause for breath, they are considered only *one* response unit if the second remark is clearly subsidiary to the first.
7. A vocal response unit may be a single word such as 'yes' or 'uh huh' or it may comprise many words such as, 'I'm going to the movies with my brother and sister and mother and father tomorrow if it doesn't rain.'
8. A single expression of affirmation ('yeah,' 'yep,' 'uh huh,' 'yes'), or of negation ('no,' 'nope,' 'nah,' 'naw'), or of interrogation ('huh,' 'what,' 'eh') may be complete responses. You are to determine by listening to the tape whether an utterance is simply a non-communicative grunt (see No. 9 below) or serves communicatively to indicate affirmation, negation, or interrogation. Examples:
   (a) do you like me (one response)
   (c) huh (one response)
   (a) I said do you like me (one response)
   (c) oh yeah (one response)
9. Expressions such as 'aw,' 'aah, 'ow,' 'haha,' 'uh,' 'oop,' when they are not used as either affirmation, negation, or interrogation do not count as responses and should be omitted from the transcripts.
10. Utterances that are not recognizable as words or word approximations do not count as responses. Examples:

(a) what color is that (one response)
(c) pa (no response)

11. Occasionally the child and adult will be talking simultaneously. For example, the adult may start to speak and the child may interject a remark so that they are both talking at the same time. If this occurs, simply separate the response of the adult from that of the child on the transcript. That is, complete typing the adult responses and then indicate the child responses on the next line.

C. *Differentiating Statements from Questions.* All responses will be marked as either statements or questions. In normal conversation questions are typically indicated by the use of particular words, by the way the words are arranged in the response, or simply by inflection.

1. Occasionally a response may start out as a question but end as a statement. When this occurs, score the response a *question.* Examples:
   (c) can I I'm going to eat my candy now
   (a) would you like me to here let me help you with that

   Both of these examples would be scored as questions.
2. A response that starts out as a statement but ends as a question is also scored a question. Examples:
   (c) I think I'll do you think it is ok to tell the aide
   (a) if I let you will you no I don't think I had better

D. *Marking the Transcripts.* You are to mark the responses in the following manner:

1. Indicate the beginning of a response by (a) underlining the first word and by (b) placing the number of the response above the first word. Number adult and child responses separately.
2. You will indicate the end of a response by placing either a single stroke (/) or a double stroke (//) after the last word.
   (a) Use the single stroke (/) when the response is a statement.
   (b) Use the double stroke (//) when the response is a question.
3. Even if the response unit consists of only one word, it is important to underline that word and follow it by the appropriate number of strokes.
4. Responses that contain words that are incomprehensible or for some other reason are excluded from the transcript will not be counted.
5. Don't forget, number adult and child responses separately.
6. It is very important that you do not fail to indicate both the beginning and ending of each response and that you number the responses accurately.

Appendix I

Prototypes of Criteria for Counting Words

1. All contractions, whether negative or affirmative, are to be considered two words (or more). Thus, contractions in expressions such as I'm, can't, won't, he's, John's talking now, etc., count as two words. Combinations such as gonna or hadda are counted as two words.
2. Expressions of affirmation (yes, yeah, uh huh), of negation (no, nope, nah, uh uh), of interrogation (what, huh), or of exclamation (oops, hey, wow) count as one word.
3. Hyphenated words and compound nouns which seem to function as single words are counted as one word each. For example:

Betty Lou	one word
Betty Lou Smith	two words
high school	one word
2-south-3	one word

4. Exclamations which tend to occur as a unit are counted as one word. For example: darn it; doggone it; oh boy; gee whiz one word each
5. Where the child is counting or is spelling, each unit (number or letter) counts as a separate word.
6. Descriptive noises such as meow-meow, grr, or bow-wow are counted as single words.

Appendix J

Instructions to Adults Regarding Form Board

Today you will be meeting individually with two of the mentally retarded children of the institution. Despite the fact that these are retarded children, it is our feeling that they can be taught to perform certain tasks if they are given adequate instruction. Our purpose today is to find out whether or not they can be taught a relatively simple but unfamiliar task if they receive individual attention. Thus, we will be evaluating the learning ability of the children as well as the teaching ability of you, the instructor.

When you finish reading this part of the instructions, you will be shown a form board which can be put together to represent a three-dimensional farm scene. You will also be provided with a diagram of the form board in its completed state. Your job will be to teach each of the children to put this device together in precisely the pattern indicated on the diagram.

In order to make the conditions standard for all children, the following rules have been established:

1) The instructional period is limited to five minutes, and you must not talk about the task or show it to the children until you are told to begin. This is very important; and since there may be some delay in setting up the equipment, you may meet some children before the instructional period begins. Should this occur, use the time to become somewhat acquainted with the child, but do not refer to the task at all. *When the session is to begin, you will hear a buzzer, and the same buzzer will sound at the end of the session.*
2) During the instructional period, you may use a duplicate of the form board and the diagram, but you may not let the child handle any of the pieces. Subsequently the child will be taken to another room where we will determine whether or not he can perform the task. It is important that the child not have handled the equipment prior to this test.
3) In order to make sure that the conditions of the experiment have been met and in order that we may watch the behavior of the children, you will be observed during the time you are with each child.

Except for the above restrictions, you are free to instruct the child in any way you deem effective. Your success will be measured in terms of how accurately and rapidly the children assemble the form board after you have seen them in comparison to how well the other children in the experiment do.

At this time I will demonstrate for you the proper assembly of the form board. Please observe carefully since this is the task that you will have to instruct the children to perform. The arrangement of objects, you will notice, corresponds exactly to the diagram, so that the animals point in a particular direction and the red apples are on the ground while the green ones are in the trees, and so on.

We believe that you are providing a good experience for the children by coming here today. We also hope that you have an enjoyable and profitable experience. Thank you for your help.

Appendix K

Instructions to Adults in Clinical Condition: Orientation Sessions

We have tape recorded a description of your job to be sure we don't neglect any important information. Because the task is a little complicated, we have arranged for you to have a couple of practice sessions today. When you come back next week, you will be asked to perform in essentially the same manner, and we will repeat the important parts of the instructions at that time. Though this may be a novel experience for you, we hope that you will find it rewarding and enjoyable.

Your main job will be to talk with each of several children. During the practice session you will talk to two children, and during the actual experiment you will meet with another four children. You will see them individually for about 15 minutes per child. We will have a timing device in an adjacent room that will buzz rather loudly when your time is up. Now I would like to tell you a little about the experiment we are conducting and about the ways in which we hope you will be able to help us.

Basically we are concerned with discovering techniques that will help some of these youngsters improve in the important area of language behavior. As the result of some of our preliminary studies of mental retardation, we have found that repeated use of questions as a device for stimulating these children is rather ineffective. Instead of encouraging children to talk, questioning by the adult seems to cause the child to 'clam up' and speak only when a specific question is directed to him. That is, while the child may answer the questions, he will usually do so in as brief a form as possible, and he will do virtually no *spontaneous* talking.

A more effective technique for getting these children to talk is for the adult to 'set the stage' for talking to occur. He can do this in several ways. One method is for the adult to engage in a fair amount of spontaneous verbalizing. If, for example, there are toys or pictures in the room the adult may 'talk to himself' about these, without requiring the child to participate. A second method is for the adult to demonstrate that he is pleased when the child does talk. He can do this by responding to the child's verbal 'leads.' Sometimes it seems to be encouraging to the child if the adult will rephrase or repeat a statement made by the child. Most people seem to enjoy hearing their ideas confirmed by others, and children are no exception.

There are also several kinds of responses that adults probably should not use in encouraging speech behavior from children. One of these, as discussed above, is direct questioning. It is also important not to apply pressure on the child or to demand that he speak. Further, the adult should avoid interrupting the child when the child is speaking. Finally, it is important that the child not be punished for speaking. That is, the child should have the feeling that the amount of talking he does and the way he does it are quite acceptable.

The question we are concerned with today is whether or not armed with these insights adults can effectively encourage retarded children to use speech

more fully and adequately. This, of course, is where you come in. We are concerned with how effectively you can carry out this program of relating to children and how effective these techniques are for increasing the child's speech proficiency.

We have provided a few simple toys which you may wish to use in helping the child become familiar and comfortable with you. You may use these at your own discretion. Each child is different, of course, and your approach to the child may have to vary accordingly. In general, however, we would hope you would be able to carry out the suggestions we discussed earlier.

There is an observation room attached to the room you will be working in. In this room we will have a recorder and we will tape the sessions. This will free you to devote your full attention to the child rather than to have to make notes during the session. Through the use of a one-way mirror it will also be possible for us occasionally to look in on the session without distracting either you or the child.

Do not get discouraged if the child does not respond to you immediately. There is a good chance that, if you use the procedures we suggested, she will eventually warm up and be more responsive, even if it is on a subsequent occasion. Now I'm going to repeat the important parts of the instructions again so listen carefully.

Today you will talk to two children. Your main job will be to see if you can't encourage them to communicate freely. You are to do this by 'setting the stage' for talking to occur. That is:

1. Give the child every opportunity to talk without pressuring or pushing him.
2. Communicate your pleasure when the child talks by being careful to respond to the child's verbalizations with enthusiasm and interest.
3. If the child is reluctant to talk spontaneously, do not ask a lot of questions; rather do some casual talking yourself, perhaps entering into some fantasy play with the toys we have provided.
4. Try to avoid direct questioning as a means of getting the children to respond to you. In the long run this will be the least effective technique.

Next week you will engage in the same procedures, but on that occasion you will see an additional four children.

Appendix L

Instructions to Adults in Interview Condition: Orientation Sessions

We have tape recorded a description of your job to be sure we don't neglect any important information. Because the task is a little complicated we have arranged for you to have a couple of practice sessions today. When you come back next week, you will be asked to perform in essentially the same manner, and we will repeat the important parts of the instructions at that time. Though this may be a novel experience for you, we hope that you will find it rewarding and enjoyable.

Your main job will be to talk with each of several children. During the practice session you will talk to two children, and during the actual experiment you will meet with another four children. You will see them individually for about 15 minutes per child. We will have a timing device in an adjacent room that will buzz rather loudly when your time is up. Now I would like to tell you a little about the experiment we are conducting and about the ways in which we hope you will be able to help us.

What we want to know is how well these children can be interviewed; that is, what kind of information can be obtained from the child about himself and about some of his experiences. We know that some youngsters at the Hospital here will be able to tell you almost nothing about themselves. We expect that you will have difficulty with some of these children. We don't know which ones; that is your job. We might mention something, however, that may help you. We have sometimes found that immediate lack of success in interviewing a child does not mean that the child can tell you nothing. You will have to work slowly with some of these children. There are not set rules for getting them to talk about themselves, for each child is different and will have to be approached accordingly. The important thing is to keep trying to work with the child right up to the last minute. Help her in whatever way you can to get her to tell you about herself; ask the child questions in any way you see fit; comfort her if necessary; give the child a piece of candy occasionally if that would seem to help her like you and want to stay with you.

We have provided a few toys in the room for the same purpose; let the child play with the toys if she wants to while you talk to her. You have been given a list of topics and will have a chance to study them before you see the first child. We don't expect you to cover every topic on the list with all of these children, just cover as much as you can in the allotted time. Some information on a topic is better than none. If you get a talkative child use the whole period to talk to her about these topics in detail. We have found that someone working with a child can do best if he or she does not have to take notes during the interview. You will need to devote your full attention completely to the child. Therefore, we have set up a microphone attached to a tape recorder in the next room to record all the things you and the child say. Through the use of a one-way mirror it will also be possible for us to look in occasionally on the inter-

view without distracting either you or the child.

Today you will talk with two children. Your main job will be to interview these children and obtain as much information from each child as you can concerning himself and his experiences. We know that some children at the Hospital here will be able to tell you almost nothing about themselves. We expect that you will have difficulty with some of the children, but we don't know which ones. We have sometimes found that immediate lack of success in interviewing a child does not mean that the child can tell you nothing. You will have to work slowly with some of these children. There are no set rules for getting them to talk about themselves; each child is different and will have to be approached accordingly. The important thing is to keep trying to work with the child right up to the last minute. Help him or her in whatever way you can to get him to tell you about himself. Ask the child questions in any way you see fit; comfort him if necessary. Give him a piece of candy occasionally if that would seem to help him or her to like you and want to stay with you. We have provided a few toys for the same purpose; let the child play with the toys if he wants to while you are talking to him. We don't expect you to cover every topic on the list with all of these children; just cover as much as you can in the allotted time. Some information on a topic is better than none. If you do get a talkative child, use the whole period to talk to him about these topics in detail.