MULTIVARIATE ANALYTIC INVESTIGATIONS FOR THE IDENTIFICATION OF SUB-POPULATIONS WITHIN APHASIA

by

Campbell McGillivrary Clark

B.Sc., McGill University, 1972

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES

Department of Education

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

August, 1978

ABSTRACT

Using medically diagnosed aphasics, this study attempted to develop a classification system based on statistical or, more precisely, variance criteria. Aphasia is a language disorder resulting from insult to the brain. Besides having direct ramifications on aphasia theory and research strategies, the techniques used here may be applicable to other atypical populations. Specifically, the outlined multivariate procedures can provide insight into intra-population distributions. The resulting distribution can then be used as a basis for factorial designs in future research.

The sample consisted of seventy-two medically diagnosed aphasics who had received speech therapy for varying time periods. Prior to and at the termination of treatment, subjects were assessed on a well-standardized and reliable aphasia battery, the Porch Index of Communication Ability. This instrument consists of eighteen subtests and purports to measure three dimensions of communication ability: one verbal, one gestural, and one graphic.

Post-treatment subtest scores were factor analyzed to determine the factor structure of the P.I.C.A. and this structure was compared with Porch's formulations. Two subtests were deleted from this analysis due to ceiling effects. Using the factor score coefficients from this analysis, factor scores for the pre-treatment subtests were generated. These factor scores were submitted to a hierarchical cluster analysis in order to determine the optimal number of groups within this sample. A series of step-wise discriminant analyses, using the factor scores as predictors, confirmed that a six-group solution was best. Group distributions were then examined with respect to current models of aphasia. The findings,

support Porch's formulations with respect to the underlying demensions of the P.I.C.A. However, the results suggested that 1) two subtests were too easy for all subjects and therefore should be discarded, and 2) two other subtests may not, in fact, reflect symbolic language function. In addition, evidence of a general language factor was also found. These findings were also compared to a factor analysis on Porch's standardization sample. The five derived dimensions of communication ability were : 1) verbal fluency, 2) writing (agraphia), 3) gestural demonstration with varying input modalities, 4) pantomime, and 5) copying. These dimensions were discussed within the context of previous aphasia research.

The grouping analysis indicated that the groups were distributed on a severity continuum not a salient features model of aphasia. Only one of the six groups was suggestive of differential impairment with respect to communication modalities, but this group was small (n = 4) and therefore must be considered judiciously.

In addition, the applicability of these procedures was discussed within the context of multivariate research. The power of multivariate designs rests on minimizing the number of groups and dependent variables and maximizing the number of subjects, and the present procedures should aid in meeting these criteria. Specifically, factor analysis reduces the number of dependent variables and yet allows for a representative sample of dependent variables. Cluster analysis groups individuals based on the similarities of their performance and thus relatively homogenous groups are formed. The distribution of these groups can form the basis for subsequent factorial designs.

W.T. Rogers ....

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

With advancing diagnostic tools and statistical techniques, empirical study of atypical populations has become both more feasible and meaningful. Methodologically, these populations present particular problems in terms of sampling, instrumentation, experimental design and subsequent analysis. In the present study, individuals diagnosed as having aphasic or language disorders resulting from insult to the brain formed the atypical population of interest. Implications of specific statistical procedures on current aphasia theories and research were investigated. First, the language dimensions underlying the disorder were examined, using a standard, reliable, and population-appropriate instrument. Second, empirical differentiation of groups within the population was investigated using dimensions of language or communication abilities derived in the first phase. By using a standardized and accepted test battery, the Porch Index of Communication Ability, the relationship between this battery and current models of aphasia was further explicated. The criteria for the development of this relationship was empirically derived rather than being nosologically or clinically based.

Etiology, Symptoms and Models of Aphasia

By definition, aphasia is a language disorder resulting from insult to the brain (Walton, 1974). The insult may be either internal (e.g. cerebral vascular accident) or external (e.g. head injury). Considerable research has been done on the recovery rates of internal versus external injured individuals. In general, the results suggest that external injuries have a better prognosis for improvement, but perhaps not for recovery, than internal injuries. The differentiation between improvement and recovery, although rarely made in the literature, may be fundamental to aphasia research. Specifically, the magnitude of the insult has direct bearing on the severity of the disorder. In other words, two 'aphasic' individuals may not appear similar in their ability to communicate or use language. For example, one may not be able to comprehend or produce language in any modality while the other may have minor word-finding difficulty when speaking. However, the more severely impaired individual has a better prognosis for improvement of language skills immediately after insult, while the less impaired individual has a better prognosis for recovery of language skills (Gerstein, 1975). This example, besides illustrating the difference between improvement and recovery, also serves to highlight the vast discrepancies found among 'aphasics'.

These discrepancies in language abilities have led researchers and clinicians to develop qualitative categorization systems based on aphasic symptoms. For example, Howes' model (Howes, 1964) differentiates aphasics dichotomously, based on verbal production from a free speech sample. The two groups are : 1) fluent - able to produce speech; and 2) dysfluent - difficulties in producing speech spontaneously. Using a psychometric bettery, Weisenburg and McBride developed a four-fold classification system as follows: 1) predominantly receptive, marked by difficulties in comprehending language; 2) predominantly expressive, marked by difficulties in speech production; 3) anomic, marked by difficulties in naming; and 4) both expressive and receptive, or mixed symptoms (Weisenburg and McBride, 1935). However, Weisenburg and McBride stressed that there was a general impairment of language function and that their model attempted distinctions based on the more predominant symptoms.

The pressure of language disorders has been associated with dysfunction in, or damage to, particular areas of the brain, usually in the left hemisphere. For example, two types of aphasia, Broca's and Wernicke's, are associated with lesions in particular areas of the brain (Walton, 1974). Generally, it is usually accepted that the left temporal lobe is associated with the decoding of auditory language units; the left occipital lobe with decoding of visual language units and the left parietal with the encoding and thus symbolic production of language units in the normal right-handed person. However, this model cannot be adhered to strictly as the brain is not clearly delineated in geographical structure and language processes are probably interactive and at times interdependent (Penfield and Roberts, 1966). The mapping of speech centres for the naming response by Penfield (1966) and more recently by Whittaker (1977), clearly indicated that a particular function may be localized but to a number of sites within the brain. Moreover, these sites may vary between and, interestingly, within individuals (Whittaker, 1977).

It is within this context of type of insult (internal versus external), predominant site and most salient presenting symptoms of insult, that aphasia research has been concentrated when examining treatment or recovery effects. For example, Weisenburg and McBride's four-fold classification system is often used to block or stratify subjects in factorial research design.

In addition, considerable effort has been directed to the development of reliable and theoretically appealing measurement instruments. Instruments in psycholinguistic models of aphasia use rating schemes for free verbal production involving such dimensions as semantic structure, correct word usage, neologism, clang associations, repetitions and amount of verbal output (Howes and Geschwind, 1964). As suggested by the dimensions assessed, these models are primarily concerned with verbal output. In contrast, psychometric batteries have been developed which attempt to evaluate an aphasic's ability to do specific tasks. These batteries may involve such skills as naming an object, demonstrating its use, verbalizing its use, writing its name, writing a sentence describing its use, writing to dictation or following verbal commands (Eisenson, 1971). Two criteria have been employed in the development of these batteries: reliability and validity with respect to a particular model of aphasia. Unfortunately, these criteria are not usually found concurrently. An obviously related question to test development *per se* is "what are the minimum number of language dimensions or tests required to describe adequately aphasic language deficits?" In a broader context, the question becomes "what is language?"

Aphasia as a Representative Atypical Population

Two factors appear to come into play in the delineation of atypical populations from the normal population: first, history or etiology and second, observable behavioural deficits or anomalies. For aphasics, the etiology may be labelled as insult to the brain, rather than a precise medical diagnosis. Within a medical context, the nature of the insult obviously has direct ramifications for treatment in terms of chemotherapy or surgical intervention. However, the more precise medical diagnosis does not have direct ramifications for behavioural intervention or therapy. Therefore, within this context, the etiology or history leading to aphasia, namely an insult to the brain, may be considered equivalent or common to all patients. The site of this insult has been associated with specific behavioural deficits. However, the initial classification rests upon impairment of language facilities. This impairment is the common behavioural deficit among aphasics. Once this initial classification is done, then further differentiation is done based on the salient features of the disorder.

The two criteria of atypical populations may become obscure due to the approaches or concerns of different areas of expertise. For example, a deaf child may be deaf for a variety of reasons (inner ear damage, or temporal lobe damage, etc.) Dependent upon the nature of this hearing loss, medical intervention may be of direct benefit. Therefore, for the medical clinicians it is of prime importance to thoroughly examine the etiology of the disorder. For the therapist attempting to retrain the child, it is of prime importance to know the behavioural nature of the loss (e.g. profound, high frequency, etc.). Therefore, two classification systems occur - one based on medical etiology and one based on behavioural deficits. However, as with aphasics, a general statement can be made, such as: "An insult to hearing apparatus has occurred which results in deficits in auditory monitoring."

The procedures to be outlined provide a means for determining intra-population stratification. As atypical populations may be comprised of relatively similar or dissimilar individuals, it would be of direct benefit to derive empirical groupings which reflect the type of behavioural deficits. Specifically, the intra-population distribution would provide estimates of population homogeneity. Aphasics, due to their etiology and behavioural deficits, are representative of atypical populations. In addition, conceptual models of deficit are well-established and thus provide a means of comparison once statistically derived groupings are established.

Methodological Problems in Examining Aphasia

Three methodological concerns arise when examining atypical populations, namely: research design, measurement and statistical analysis. One area of enquiry which has been overlooked in aphasia research is how well this population conforms to the basic tenets underlying empirical or experimental investigation. It is the intent of this study to show that, by investigating this question, further insight may be gained into the phenomena of aphasia, as well as into the development of better research methods for its study. In order to investigate this question, the classical experimental model as espoused by Fisher (1935) will first be reviewed briefly and then the phenomenon of aphasia will be investigated in the context of this experimental model.

Fisher suggests that the population of interest be randomly sampled, selected members randomly assigned to groups, treatment and non-treatment applied, and then the groups compared on a previously defined variable(s), in order to determine if differences exist on these variables which can be ascribed to differential treatment effects. Certain assumptions are made with respect to the population of interest, the treatment(s) and the dimensions affected by treatment, so that statistical investigations can be performed. Specifically, it is assumed that the population of interest is the universe and that the dimension(s) of interest is (are) distributed normally and independently within the universe. Therefore, by randomly sampling the population of interest, a group or groups representative of the population may be obtained such that: 1) error can be assumed to be randomly, normally and independently distributed within the group(s); and 2) the assumptions of homogeneity of variance in the univariate case and homogeneity of variance and covariance in the multivariate case are not violated. Treatment is then applied, implying it is strictly defined and controlled. This process then allows for the treated and untreated groups to be compared with the assumption that variability may be partitioned into that which is related to treatment and that which related to individual differences within the population (i.e. error variance). Given the aforementioned assumption, this error variance is then distributed as χ^2 with the appropriate number of degrees of freedom; similarly, under the null hypothesis, the treatment variance is assumed to be distributed as χ^2 also and thus gives rise to a central F-value. However, if the treatment variance is not distributed in accordance with the error variance - in other words, as a non-central F-value - then the researcher fails to accept the null hypothesis (Morrison, 1975).

When the phenomenon of aphasia is examined within the context of this experimental model, certain questions arise based on etiological concerns and the definition of language dimensions or function. First, can language dimensions be assumed to be normally and independently distributed within the aphasic population? In other words, are language dimensions independent or dependent and how are they distributed within a population which has arisen due to untoward and uncontrolled environmental events which result in insult to the brain? In order to examine these questions, one must assume that the initial treatment is the insult to the brain. This treatment, by its nature, is uncontrolled and uncontrollable. Specifically, the treatment is uncontrolled for in terms of site, severity and type of insult. The assumption that insult to brain causes marked and measurable changes in an individual's ability to perform certain or many tasks has been well-established by case studies, normal versus head-injured designs, and clinical observation (Walker et al, 1969; Klonoff et al, 1978). From these forms of research a higher order question emerges: differential groups arise from insult to the brain? Given the nature of brain insult, it is morally and ethically impossible to 'control' treatment; rather, one must examine existing groups post trauma. This change in the temporal ordering of Fisher's experimental paradigm -- from sampling, treatment, then assessment of difference to (uncontrolled) treatment dictating selection -- forces the researcher to use alternate approaches for the evaluation of differential changes subsequent to trauma. Unless the meaningful aphasic groupings can be delineated subsequent to trauma and before remedial treatment, the evaluation of post trauma treatment modalities is confounded at best. Therefore, post hoc attempts to classify aphasic symptoms after trauma but before remedial treatment using statistical -- or, more specifically, variance/variability considerations -- would seem warranted.

As mentioned previously, the phenomenon of differential effects has long been recognized in aphasiology. However, there is evidence to suggest that the models suggesting differential classification system may be unduly influenced by historical events affected by the type, site and severity of insults (Golden, 1977). In addition, the development of these classifications systems was not related to the systematic or empirical partitioning of variability or variance based on measured test performance. For example, Howes dichotomous classification system is based on a differential effect model, and uses only verbal production to discriminate groups, while Scheulle's (1969) model is based on a severity continuum. Neither researcher has systematically examined his data to determine the optimal partitioning of variability in terms of group membership (i.e. number and type). In addition, both researchers highlight just one aspect of aphasia. However, recent developments in computer technology now make possible the examination of performance in terms of multivariate models.

The importance of such procedures becomes apparent when the question of treatment efficacy is considered. In particular, if insult to brain causes different types of aphasic disorders, then it is reasonable to hypothesize that different treatment strategies will have differential effects on specific groups. Even within a severity model such a procedure is relevant because the level of severity dictates the level of accessibility to specific treatment modalities. Therefore, in order to generate more powerful designs to gauge the efficacy of specific treatments on particular types of aphasics, an examination of the characteristics of an aphasic population, and how these characteristics are distributed, would seem warranted. Historically, such an approach has been used by Head (1926) and Weisenburg and McBride (1935) -- without, however, the advantage of current multivariate statistical analysis techniques.

Given that language disorders have many dimensions or possible areas of deficit, a multidimensional test battery must be used in order to assess fully the severity and type of disorder. The multidimensional aphasia test battery must also be reliable and a valid measure of communication abilities. Most instruments presently available are screening tests using dichotomous scoring techniques, such as the Halstead Wepman Aphasia Screening Test (1959) and designed on the basis of an existing model of aphasia such as Eisenson's Examining for Aphasia (1954). The standardization of many aphasia tests has been poor due to the limited number of subjects available within a given setting. One test battery, containing 18 subtests, which was designed and standardized to remedy this situation is the Porch Index of Communication Ability (P.I.C.A.) (Porch, 1968). Particular attention was paid to development of a reliable and thereby generalizable instrument in terms of communication deficits reflected by a particular set of scores. In addition, the tasks selected were common and recognized tests of aphasia deficits. Therefore, the instrument of choice for this study was the P.I.C.A.

Multivariate Statistical Concerns and Procedures Relevant to the P.I.C.A.

Although aphasic disorders have not been viewed conceptually as a unidimensional problem, technology has limited analysis strategies. Rather, it has been suggested that language disorders are multidimensional, and to assess fairly a language disorder multiple tests should be given in order to determine areas of asset and deficit. Such a model clearly suggests that multivariate statistical techniques are required as well as multidimensional assessment instruments. However, as the power and sensitivity of multivariate designs are determined by the number of subjects, groups and dependent variables, it is essential to minimize the number of dependent variables and groups based on variance considerations, while maximizing the number of subjects (Morrison, 1975). The purpose of this study was first to determine the optimum number of language dimensions to describe aphasic performance elicited by the P.I.C.A. by factor analysis and then to determine what the optimum grouping of subjects would be based on these dimensions by hierarchal grouping analysis. Such a procedure would then allow for group profiles based on a finite number of language dimensions to be produced. These profiles would then allow the salient characteristics of group membership to be determined.

Factor Analysis

As suggested by Porch and confirmed by factor analysis (Clark et al, in press), there are three dimensions of language abilities underlying the P.I.C.A. Therefore, it would be redundant to include all eighteen subtests in a multivariate design. This redundancy makes the explanations cumbersome, particularly in view of the limited sample size readily available for research work. For example, in a two group design, eighteen degrees of freedom would be used by dependent variables rather than perhaps three. The reduction in degrees of freedom has two major advantages. First, degrees of freedom which would have been ascribed to the dependent variables would now be associated with the error variance thus reducing the error variance estimate. Second, as the number of overall tests of significance is reduced, the overall experiment-wise error is also reduced. Therefore, a factor analysis of the data may significantly reduce the number of dependent variables as well as providing better insight into the true dimensions underlying the variables and thus in the development of a clearer conceptual model (Gorsuch, 1974). For this study, the first analysis was a factor analysis of data obtained on the eighteen P.I.C.A. subtests. The language dimensions derived from this analysis were then compared to both Porch's formulation and other aphasia models of communication ability.

Grouping Procedures (Hierarchical Cluster Analysis)

Pre treatment factor scores were then derived and submitted to a hierarchical cluster analysis in order to determine the minimum number of groups required to adequately represent subsamples within the sample (Patterson and Whittaker, 1973). The criteria for this grouping decision was variability considerations in terms of minimizing within group variability, while maximizing between group variability. Although this solution will optimize groupings, it should provide insight into possible factorial strategies in future research. To investigate how well these groups were differentiated by this procedure, a step-wise discriminant analysis was done on possible solutions (Nie et al, 1975). Although the prediction rates and F-values were artificially inflated, a rough indication of group homogeneity was obtained. In addition, the group profiles were plotted, compared with each other and with existing models of aphasia.

Summary of Procedures and Their Conceptual Ramifications

In order to do this study, a group of aphasics (n = 72) were assessed on a multidimensional aphasia test battery, the P.I.C.A. The sampling of aphasics was not stratified on an existing model of aphasia. Rather, the sample consisted of consecutive admissions to a rehabilitation hospital and to whom speech therapy for aphasia was given. These aphasics were given the P.I.C.A. both pre and post- treatment. The post-treatment subtest scores were then factor analyzed and factor scores for the pre-treatment scores were estimated. These factor scores were then submitted to a cluster analysis. The derived grouping were then examined with respect to their distribution in vultivariate space.

These procedures allowed for: 1) the language dimensions underlying aphasic disorders to be examined; and 2) the empirical differentiation of groups within this population.

CHAPTER II

LITERATURE REVIEW

Introduction

Three main topics pertinent to aphasia are reviewed in this chapter. First, historical and current models of aphasia are discussed. Second, methodological advances are briefly described, with particular reference to multivariate approaches. Finally, the standardization and psychometric properties of the Porch Index of Communication Ability (P.I.C.A.) are reviewed.

Models of Aphasia

The phenomenon of aphasia has been recognized since biblical times and explanatory constructs have been based on grounds ranging from philosophical to medical. Interestingly, the loss of speech or language and the associated right-sided motor involvement was considered God's retribution for sin in earlier times. Although today's models of aphasia are somewhat more advanced in terms of causation, they are not appreciably more advanced in terms of facilitating recovery or prognosis. Within conceptual models of aphasia, three separate approaches to the problem may be differentiated as follows: 1) neurological - based on the site of the lesion; 2) behavioural - based on salient features of deficits; and 3) linguistic - based on deficits in language usage, grammer and facility. Although in practice these models are not independent, for the purpose of explication they are treated separately in this review.

Within a strict model of neurology, the diagnosis of aphasia is only made when the diagnostic data indicates a manifest disruption of left hemisphere function and this disruption can be shown to affect language facility adversely (Walton, 1974). With the advancing diagnostic techniques (e.g. C.A.T. and Brain Scans, E.E.G.'s), the identification and localization of these focal or diffuse disruptions have become more feasible and accurate. For example, Naeser and Hayward (1977) have shown, using computerized axial tomography (C.A.T.), that the severity of the aphasia disorder is directly related to the size and density of the blood mass or swelling of the blood vessels. For cases of cerebral vascular accidents resulting from embolis, aneuryisms or thrombosis, such a procedure provides insight into both the magnitude of the disorder and by serial scans, the amount of recovery or degeneration of brain tissues. In this model, the differentiation of dysarthria, the inability to produce speech due to involvement of neuromuscular mechanisms, from aphasia is based on the site of the lesion. Specifically, for the diagnosis of dysarthria to be made, the lesion must occur in the brain stem.

Three objections to this model occur when other criteria are considered. First, the patient may not be able to produce speech with a lesion in the left hemisphere, but tests of language using other modalities (e.g. writing) suggest language function is preserved. Therefore, within the confines of the neurological diagnostic model, the patient is aphasic, but within a speech/language continuum, the patient is dysarthric. Second, for the diagnosis of aphasia to be made, the lesion site must be confirmed to be in the left hemisphere and the patient must exhibit speech/language impairment. The question arises concerning what is language impairment, in that perfect language or normal language facility has not been clearly defined (Lesak, 1976). For example, in a mild conduction-type aphasia, the patient can understand commands and accurately respond in a verbal modality to historical questions, but has great difficulty in generating spontaneous speech. On specific language tests, the degree of deficit may be profound but, within the normal social context, not noticeable. Rather, the patient may be described as socially retiring or shy (Goodglass and Kaplan, 1972). The third objection to this model is that it assumes brain-behaviour relationships are invariate. Namely, language/speech mechanisms are lateralized to the left hemisphere. Although a relationship does exist, the evidence for exceptions to the model is large. The major population which contradicts or at least does not wholly support this relationship is left-handed individuals, where in language/speech functions may be lateralized to either hemisphere or distributed between the hemispheres (Lezak, 1976).

Behavioural models of aphasia attempt to categorize aphasics based on the most salient clinical features. Two historical types of aphasia which are still recognized diagnostic categories are Broca's (1861) and Wernicke's (1881). Both these types of aphasia were related to specific lesion sites in the left hemisphere found on autopsy. In addition, both have characteristic clinical or behavioural features. Broca's aphasia is marked by the inability to produce meaningful speech while the patient still comprehends verbal input. Wernicke's is marked by the inability to comprehend speech or language while verbal production is relatively intact. With these types of aphasia, a site-function school of brain function developed. It was felt that each behavioural action or, in these cases, deficit was related to a specific and local dysfunction in the brain. This premise led to the development of behavioural classification systems based on the most salient clinical features.

It should be noted, however, in contrast to the site-behaviour schools of brain damage, Hughling Jackson and his followers espoused that insült to the brain caused general dysfunction of cognitive process and although some deficits were more severe, cognition per se was decreased. Jackson further suggested that this general dysfunction was of more importance than specific deficits with respect to a patient's day-to-day function. However, this view was considered the minority (Head, 1926).

Modern behavioural theories of aphasia still reflect this controversy of general versus specific deficits related to brain damage. Two recent proponents of the generalist position are Luria and Milner (Luria, 1966; Milner, 1972). Both based their positions on a language or communication model and argued that insult to any part of the cortex adversely affects one's ability to communicate. For example, if a patient has a deficit in spatial planning (hypothesized to be right hemisphere function) then the patient's ability to describe or orally interpret this type of problem will be also reduced. Interestingly, besides their own research, current research also support this position. Winner and Gardner (1978) found that right hemisphere damaged patients had more difficulty in interpreting metaphora and tended to choose literal interpretations of metaphors, while patients with left hemisphere anterior lesions (aphasics) approximated normals in their performance and seemingly still understood metaphorical language. Both Luria's and Milner's positions argue for an interactive and interdependent model of brain damage. Both suggest that using a deficit or site of lesion with associated specific cognitive losses model is too simplistic.

However, the majority of current categorization systems assume that aphasia is solely a language problem. More importantly, these systems maintain that observable or measurable deficits in other forms of cognition are purely a function of a patient's deficits in language facility. Schuelle's (1969) approach probably represents the bridge point in that it supports the first assertion, but disputes the second. Scheulle contended that the major difference among aphasics is the severity of the impairment, not specific loss of function (e.g. auditory comprehension). Thus, aphasia is a unitary concept representing language impairment. However, she also contended that aphasics may have other lesions affecting visual perceptual skills (i.e. right hemisphere function) or motor function (i.e. brain stem). Accordingly, she generated the following five-fold classification system: 1) simple aphasia; 2) aphasia with visual perceptual deficits; 3) aphasia with motor involvement; 4) aphasia with visual perceptual and motor deficits; and 5) inaccessible to examination.

Weisenberg and McBride's (1935) model of classification is based solely on a language model and does not involve other cognition systems. Their research also supported a general language impairment model, but also suggested that patients could be differentiated further on a salient feature model. Their groupings were as follows: 1) predominantly expressive (Broca's); 2) predominantly receptive (Wernicke's); 3) both receptive and expressive deficits; and 4) anomic, marked by word finding difficulties, but with fluency and comprehension relatively intact. Although Weisenberg and McBride did stress that there was general language impairment, current clinical and research practices tend to ignore the general aspect of impairment in favour of highlighting the salient features. Rarely does one encounter research where subjects are first differentiated on severity continuum, then on a salient feature classification. Rather, the factoring or blocking of subjects is based on an expressive/receptive dichotomy. In a clinical setting, the diagnosis of, for example, expressive aphasia is misleading or at least incomplete because auditory comprehension may also be reduced.

Probably, the most comprehensive behavioural model of aphasia has been suggested by Wepman and Jones (1961). Using the expressive/receptive dichotomy, Wepman and Jones further expanded it through all input/output modalities. Their classification system is too extensive to review here, but it will suffice to say that all input modalities -- e.g. reading of letters, words, sentences, aural recognition and comprehension of letters, words and sentences and tactile recognition -- and all output modalities -- e.g. writing, speech and gestural -- are evaluated. They argued that behavioural description of specific deficits is the best means of classifying aphasics.

One major departure from traditional psychometric evaluation of aphasics was done by Howes (1964). Howes viewed aphasia as primarily a speech/language problem and felt traditional psychometric assessment involved many higher level constructs (e.g. intelligence) not necessarily related to aphasia per se. He therefore designed a structured method for gathering speech sample. Analysis of this sample using production variables -- such as the number of words, pauses, and meaningless utterances -- suggested a dichotomous model of aphasia. Howes labelled his two groups fluent or dysfluent. It has been suggested that these groupings are consistent with Weisenburg and McBride's expressive/receptive classification system (Howes, 1964). Therefore, by using a standardized but different assessment technique, a type of cross-validation has occurred if one examines the differences in labelling in the respective systems.

More importantly, however, Howes' method interested psycholinguists and has led to psycholinguistic models of aphasia based on language development or specific linguistic deficits. These models are beyond the scope of the measurement procedure adopted in the present study and therefore will not be reviewed here. However, it should be noted that psycholinguistic models of aphasia have grown in popularity and represent a measurable proportion of current research studies (Williams, 1976; Zurif et al, 1972; Zurif et al, 1974; Goodglass, 1975).

All these classification systems were developed using standard and specially designed psychometric testing techniques. The problems encountered in testing aphasics, scoring responses and integrating results have been well-reviewed (Osgood and Miron, 1963) and represent a volume in themselves. However, biases and assumptions within the tests themselves obviously influence interpretations and subsequent theorizing. Factor-analytic studies on these test batteries usually report a large first eigen-value, thus suggesting a general language component. However, specific dimensions also keep occurring across different test batteries (Carroll, 1941; Scheulle et al, 1962; Clark et al, in press). Moreover, modern analysis techniques and computer facilities are only now allowing researchers to gauge the relative weights of these dimensions with respect to a severity model. Interestingly, the recent articles support these historical models, as well as reflecting the complexity and subtleties of the problem.

History of Research Methods

The history of methods employed in aphasia research is consistent with the development of methodological standards, only delayed in terms of time. The phenomenon was well established and recognized by the turn of the century. In addition, causation paradigms based on concrete evidence were being espoused, e.g. those of Broca and Wernicke. In 1921, Head recognized the need for an aphasia test battery and developed one based on the assessment of sixty-seven aphasics. Using a case study methodology, Head developed probably the first comprehensive battery for the assessment of aphasia. In 1932, Weisenburg and McBride realized the importance of a control group in differentiating real aphasic deficits from individual differences in normal language abilities. In their work, they used Head's original tests as well as other standardized psychometric batteries in an attempt to develop a test battery for distinguishing aphasics from the general population.

In 1941, Carroll realized the power of factor analysis in delineating the underlying dimensions of aphasia. Although limited by the available technology, Carroll's work represents the first multivariate analysis attempt on aphasic populations. Until this time, research analyses were confined at best to recognized univariate procedures.

During the 1950's, aphasia research centred primarily on test development and classification procedures. Factor analysis played a major role in delineating language dimenisions. However, groupings systems were then based on derived factor dimensions, not upon the discrimination function amount groups using these dimensions (Schuelle et al, 1962). The tests developed were designed primarily to reflect existing classification systems. For example, Eisenson's Examining for Aphasia, reflected Weisenburg and McBride's classification system (Eisenson, 1972). However, dimensions underlying items were assumed to be receptive or expressive skills, but these assumptions were not impirically tested.

In contrast to these 'statistically' based models or examinations of aphasia, the Boston School reverted to a case study approach very similar to that of Head's. At present, they recognize nineteen discrete types of aphasia. Again, this model is based on salient characteristics rather than severity. Moreover, their model has sixteen dimensions of aphasic language (Goodglass and Geschwind, 1975). Obviously, some of these dimensions must be highly inter-related (e.g. articulation and speech fluency) or their categorization system would be even more extensive than it is now (i.e. 2^{16} -- if the dimensions were truly dichotomous and independent).

Multivariate statistical research on the relationship between language dimension and subsequent grouping procedures is extremely sparse. For example, only one study could be found which used both factor and cluster analysis. Crockett (1976) both factor analyzed data obtained on the Neurosensory Centre Comprehensive Examination of Aphasia (NCCEA) and derived groupings based on cluster analysis. However, these procedures were applied independently on the eighteen variables of the NCCEA. These procedures were also applied to a free speech sample measured over sixteen variables. As the purpose of the study was to show points of interface and difference between these instrumentation procedures, the analyses were appropriate. Similarly, the study examined continuity of fit between statistically derived groupings and the two aphasia models represented by the instruments (Weisenberg and McBride; Howes and Geschwind). In this sense, the study did not examine naturally occurring groups based on derived factor dimensions; rather it examined the relationship between performance and two existing models of aphasia.

The P.I.C.A. as an Assessment Instrument of Aphasic Disorders

When Porch developed the P.I.C.A., it was with the expressed purpose of producing a multidimensional battery of language function. The battery consists of eighteen subtests with ten items per subtest. The tasks demands within each subtest are the same (e.g. naming ten different objects). The task demands for each of the eighteen subtests are summarized in Table 1. Porch suggested that these eighteen subtests measured three major language/communication dimensions: verbal, gestural and graphic. The presence of these three language/communication dimensions was recently found in a factor analysis of Porch's reported subtest intercorrelation matrix (Clark et al, in press).

For a scoring system, Porch developed a multidimensional model in which the quality of the elicited response to each task is quantified. Five dimensions of response quality were subsumed in the scoring system: 1) accuracy -- the degree of correctness; 2) responsiveness -- the ease with which the response is elicited; 3) completeness -- the degree to which the task is carried out in its entirety; 4) promptness -- response latency; and 5) efficiency -- the degree of facility demonstrated in performing the motoric aspects of the task. The scoring range is from 1 (no response) to 16 (a complete, prompt, and accurate response). A schematic overview of this scoring system is presented in Table 2 and Figure 1. During standardization, twelve clinical judges were asked to rank these descriptions of response quality in terms of severity. The coefficient of concordance for these judgements was 0.93.

TABLE 1

Description of P.I.C.A. Subtests

Subtest	Response Modality	Task for Subject
I	Verbal	describe purpose or function of object
II	Gestural	demonstrate function of object (ordered)
III	Gestural	demonstrate function of object (unordered)
IV	Verbal	name object
V	Gestural	read card and carries out instructions (functional description)
VI	Gestural	identify object based on functional description
VII	Gestural	read card and carries out instructions (naming descriptions)
VIII	Gestural	match picture card with object
IX	Verbal	complete sentence by naming object described
X	Gestural	point to object named
XI	Gestural	match similar objects
XII	Verbal	repeat name of object after examiner
A	Graphic	write sentence describing use of object
B	Graphic	write names of objects
C	Graphic	write name of object to auditory dictation
D	Graphic	write name after having it spelled
E	Graphic	copy name
F	Graphic	copy geometric shapes

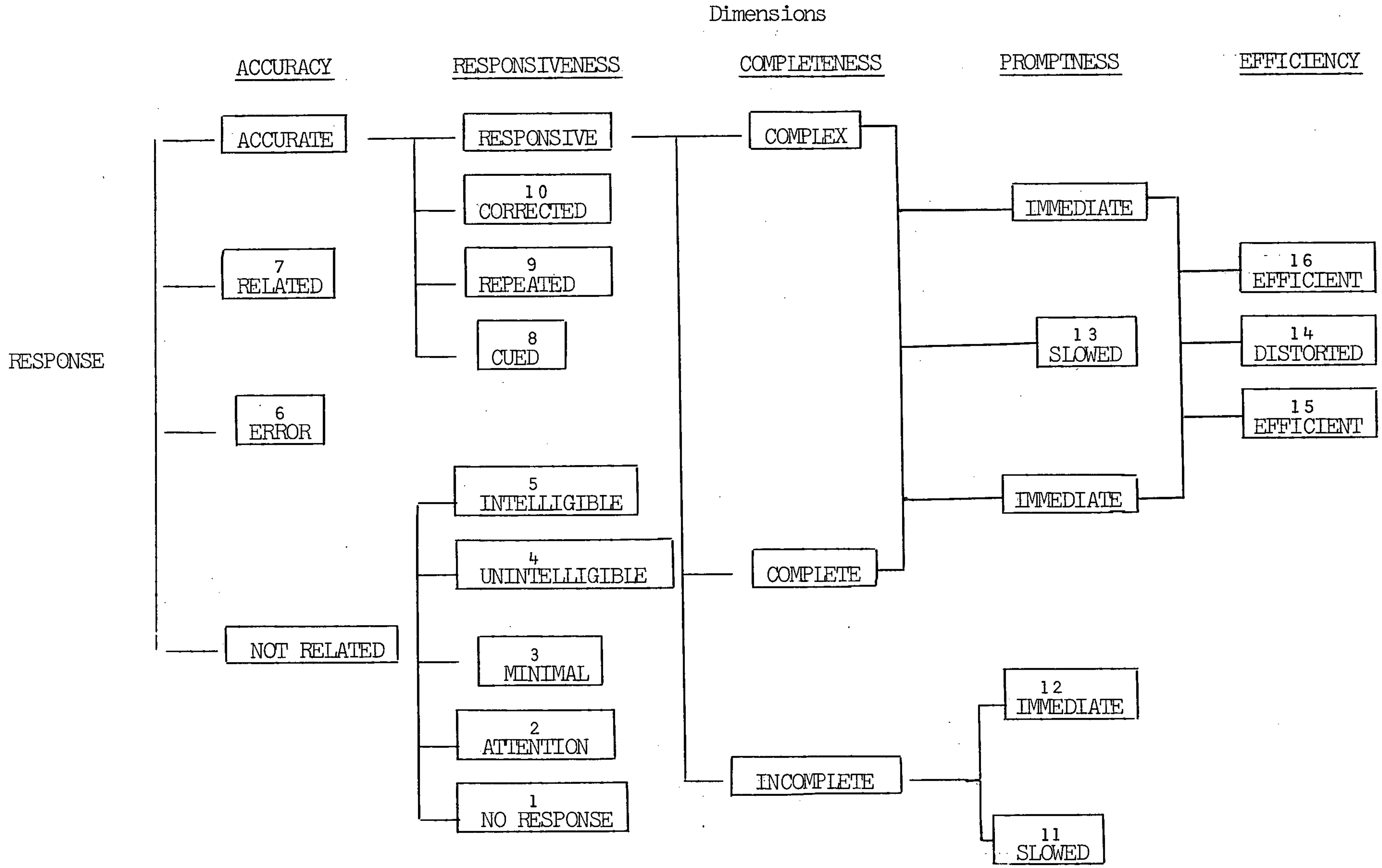
TABLE 2

Multidimensional Scoring Categories of the P.I.C.A.

Score	Level	Description
16	COMPLEX	Accurate, responsive, complex, immediate, elaborative response to test item.
15	COMPLETE	Accurate, responsive, complete, immediate response to test item.
14	DISTORTED	Accurate, responsive, complete response to test item, but with reduced facility of production.
13	COMPLETE-DELAYED	Accurate, responsive, complete response to the test item, which is significantly slow or delayed
12	INCOMPLETE	Accurate, responsive, response to test item which is lacking in completeness.
11	INCOMPLETE-DELAYED	Accurate, responsive, incomplete response to test item which is significantly slowed or delayed.
10	CORRECTED	Accurate response to test item self-correcting a previous error without intervening variables.
9	REPETITION	Accurate response to test item after a repetition of the instructions by request or after a prolonged delay.
8	CUED	Accurate response to test item stimulated by a cue, additional information, or another test item.
7	RELATED	Inaccurate response to test item which is clearly related to or suggestive of an accurate response.
6	ERROR	Inaccurate response to the test item.
5	INTELLIGIBLE	Intelligible response which is not associated with the test item, e.g., perserverative or automatic responses or an expressed indication of inability to respond.
4	UNINTELLIGIBLE	Unintelligible or incomprehensible responses which can be differentiated from other responses.
3	MINIMAL	Unintelligible response which cannot be differentiated from other responses.
2	ATTENTION	Patient attends to test item but gives no responses.
1	NO RESPONSE	Patient exhibits no awareness of test item.

FIGURE 1

Porch's Multidimensional Scoring System



BINARY CHOICE SCORING SYSTEM BASED ON FIVE DIMENSIONS

The battery was standardized on 150 medically diagnosed aphasic patients. No attempt was made to apply a stratified sampling technique based on existing conceptual models of aphasia. However, in comparison to other tests of aphasia, this battery has a large standardization sample. Further, it represents a departure from other tests which consist of tasks which differentiate normals from aphasics. Due to the nature of the tasks on the P.I.C.A., any normal speaking person should obtain a perfect score.

With respect to psychometric reliability criteria, this test is extremely impressive. Internal consistency estimates for the eighteen subtests range between 0.88 and 0.99. These results would suggest that the multidimensional socring system is consistent across items for each subtest. Moreover, it supports the conceptual model that the task demands of specific subtests are similar across items. Reliability coefficients across raters are of the same magnitude across the eighteen subtests -- 0.93 to 0.99. These findings would suggest that if raters are trained in the admission of the P.I.C.A., score assigned by different raters would be extremely similar. Stability coefficients on a test-retest model with a two week interval range from 0.70 to .98 for the eighteen subtests. However, only one subtest's coefficient was below 0.81. These results would indicate that the dimensions of language ability the subtests purport to measure are stable over time. In addition, the eighteen subtests appear to comprise a difficulty gradient, as mean scores range from 6.01 to 14.22. These reliability estimates, as well as the mean for each subtest, are presented in Table 3.

TABLE 3

Measures of Reliability for the P.I.C.A.

Tests	In Order of Difficulty Means†	Inter-Rater Reliability	Internal Consistency Hoyt's Anova	Stability Test-Retest(<2 weeks)
A	6.01	0.93	0.97	0.95
B	6.62	0.97	0.97+	0.92
C	6.94	0.97*	0.96+	0.96
D	7.39	0.97	0.96+	0.92
I	8.40	0.99	0.98	0.97
E	8.56	0.94*	0.95+	0.85
V	8.87	0.99	0.99	0.90
II	9.37	0.95	0.93+	0.87
IV	9.45	0.99	0.98	0.97
IX	9.48	0.98	0.98	0.96
VII	9.94	0.99	0.99	0.91
III	10.36	0.93*	0.95+	0.88
F	10.49	0.95	0.96+	0.98
XII	11.17	0.99	0.99	0.95
VI	12.00	0.99	0.96	0.94
X	12.67	0.99	0.96+	0.81
VIII	13.44	0.98	0.96	0.95
XI	14.22	0.95	0.88+	0.70
n	150	30	150	40

* significant difference ($p < .05$) found between mean scores of raters.

\+ significant difference found between certain scores of subtest items.

\+ standard deviations not given in manual.

As mentioned earlier, a factor analysis of Porch's reported subtest interocrrelation matrix supported the presence of the three language dimensions posited by Porch (Clark, et al, 1978). As the specific subtests loaded, in the main, on factors consistent with Porch's formulations, this factor analysis provides evidence to support the construct validity underlying the P.I.C.A. Due to the magnitude of the first eigenvalue obtained from a principal component analysis on the reported inter-correlation matrix and the high and reasonably equivalent inter-correlations among factors suggested by the optimal oblique transformation, a second order factor analysis was also performed. The results of this analysis suggested that a higher order factor of general communication abilities was also present. Interestingly subtest loading on this factor was positively correlated with the difficulty of each subtest (Spearman's $r = 0.68$). Specifically, the easier tests were more highly correlated with the general language factor. Therefore, this analysis gave support to both severity and specific deficit theories of aphasia. The factor pattern and structure matrices are given in Appendix A, Tables 1 and 2.

The reported findings on the P.I.C.A. as a psychometric instrument, as well as the standardization procedure employed, indicate that it is an appropriate test battery to evaluate aphasics and to explore the dimensions underlying this disorder.

CHAPTER III

METHODS

Introduction

In the first section of this chapter is presented an overview of the experimental paradigm and an outline of the statistical procedures employed. In addition, this section also identifies the areas of conceptual interest to be examined later for specific stages of the analysis. The second section of this chapter presents the criteria for the specific types of analysis employed. As many of these analyses either violate traditional assumptions or capitalize on certain aspects of the data, the rationale for their use is also provided.

Experimental Paradigm and Statistical Procedures

The temporal ordering of treatments and evaluation of aphasic symptoms are presented in Figure 2. As mentioned earlier, it has been shown that aphasics are different from the normal population. However, the distribution of skills or deficits at time 0_1 has never been examined using multivariate statistical techniques. Previous attempts to classify or group subjects used univariate techniques (Howes, 1964), neurological models of deficit or expressive/reception dimensions (Weisenburg and McBride, 1935). As statistical analysis represents the main form of argument in terms of treatment efficacy for groups, this study attempted to explore potential groupings using statistical criteria. The examination of possible groups post-trauma but pre-treatment (i.e. 0_1) would then provide a factorial framework for subsequent research to examine differential treatment effects.

FIGURE 2

Overview of Experimental Paradigm: Aphasic Research

Time:	X_1	O_1	X X	O_2
	↓	↓	↓	↓
	'treatment' resulting in aphasia uncontrolled, environmental event	assessment of changes subsequent to trauma	treatment leading to recovery/ improvement e.g. time, relearning of skills, development of coping strategies	assessment of residual or pure aphasic symptoms

TABLE 4

Data Analysis Procedure

STEP 1. Calculation of mean and standard deviation for eighteen subtests at 0_1 and 0_2.

2. Examination of mean and standard deviation to determine suitability for Factor Analysis.

3. Factor Analysis of subtests retained at 0_2.

4. The generation of factor scores for subtests scores at 0_1 using the derived coefficients computed at 0_2.

5. Hierarchical grouping analysis of factor scores at 0_1.

6. Discriminant analysis of possible solutions to gauge the efficacy of specific grouping procedures.

7. Calculation of means and standard deviations for optimal solution.

8. One way analyses of variance by group on derived factor dimensions.

9. Comparison of group means using Scheffe's multiple comparison techniques.

10. Qualitative statements concerning specific factor dimensions and distribution of groups.

11. Qualitative statements concerning group profiles.

With respect to statistical methodology and techniques, the proposed procedures are presented in Table 4. Steps 1 and 2 represent the examination of raw data to ascertain its suitability for factor analysis and allowed for comparisons between this sample and Porch's standardization sample. The factor analysis (Step 3) of the subtest scores at 0_2 reduced the number of dependent variables. In addition, the derived factor pattern was compared to Porch's initial formulations and the factor analysis of the standardization sample (Clark et al, in press). In Step 4, factor scores were generated for the pre-treatment assessment using the coefficients derived from the factor analysis at 0_2. Steps 5 through 7 involved evaluation of the grouping procedure. Steps 8 and 9, although allowing spurious results in terms of classical statistical theory, produced an estimate of group differences. Moreover, by using the most conservative multiple comparison technique available, Scheffe's, the number of reported differences were minimized, not maximized. By examining each factor separately, the distribution of groups relative to current aphasia theory was also reviewed. Finally, the group profiles were examined.

<u>Sample</u>

The sample consisted of seventy-two medically diagnosed aphasics who had been admitted to the G.F. Strong Rehabilitation Centre for speech therapy. The criteria for inclusion in the sample were: 2) therapy must have lasted at least three months, unless the patient was discharged because speech therapy was no longer deemed necessary; b) P.I.C.A. examinations prior and subsequent to treatment must have been done on all subjects; and c) all protocols must have been complete (eighteen subtests). Demographic data were collected in terms of age, sex, education and diagnosis.

Instrumentation

The psychometric properties of the P.I.C.A. have been reported in Chapter 2. Subtests on which a ceiling effect was observed ($\overline{X} > 14.0$) were discarded. Inclusion of such subtests would affect the factor analysis adversely. Specifically, factor dimensions could arise which reflect variables with no variance. In addition, the mean scores attained by this sample were compared to those reported by Porch. As Porch reported no standard deviations, a Hoteling T^2 one sample procedure was used with Porch's means considered as parameter estimates.

ANALYSES

Factor Analysis

All subtests which met the criterion for inclusion were then factor analyzed. The post-treatment subtest scores (i.e. 0_2) were used for this analysis as these scores best reflect residual aphasic disorders after treatment and thus more accurately reflect true aphasic disorders. The factor analysis was completed in two steps. First, possible solutions were determined based on the configuration of the eigenvalues derived from a principal component analysis at 0_2. An initial principal components extraction was obtained using the Alberta General Factor Analysis Programme (Hakstian and Bay, 1972). The resulting eigenvalues were examined to determine the possible number of factors to retain. Three criteria were considered: 1) the number of eigenvalues greater than one (Kaiser-Guttman Rule) (Guttman, 1954); 2) the determination of inflection points from the graph of eigenvalues versus ordinal number of factors (Scree Test) (Cattel, 1966); and 3) the percentage of accounted variance. For the second stage of the analysis, factor solutions were then obtained for each of the number factors suggested by the aforementioned criteria using unweighted least squares analyses (Hakstian and Bay, 1972). The residual correlation matrices of each solution were examined and the number of correlations less than 0.15 counted. As the major purpose of this study was exploratory, an overfactored solution was considered better than an underfactored solution.

The preferred solution was then transformed using a series of the Harris-Kaiser Method of Oblique/Orthogonal Transformations. The best solution was determined by conforming to simple structure: the factor pattern with the highest number of loadings <0.15 (hyperplane count) and with the fewest number of variables loading >0.30 on more than one factor (factorial complexity). In the case where solutions were equivalent in terms of these criteria, the less oblique solution was selected.

In order to generate factor scores, the factor score coefficient matrix was then standardized by pre-multiplying it by the inverse diagonal overall standard deviation matrix. The resulting matrix was used to estimate pre-factor scores by the regression method (Hakstian and Bay, 1972).

Hierarchical Analysis

The pre-factor scores were then submitted to a hierarchical grouping analysis in order to determine the 'correct' or optimal number of groups. This procedure involves both minimizing within-group variability and maximizing between-group variability (Patterson and Whittaker, 1973). Initially, the analysis began with the number of groups equal to the number of subjects (i.e., no within-group variability, maximal between-group variability). Similar subjects were then collapsed into a group at each step until there was just one group solution (i.e. maximal within-group variability, no between-group variability). At each intermediate step, the cumulative within-group variability term, the change in this term from the preceding step, and a stability index (which relates the change in the error term to overall variability and corrects this ratio by the number of groups) were calculated. In addition, the percentage of within-group to the overall- variability, and the change in this percentage, were calculated. To determine the ideal number of groups for maximizing between-group variability and thus minimize within-group variability, the following criteria were considered: 1) percentage of overall variance accounted for between groups; 2) the relative change in this percentage as the number of groups decreased; and 3) the stability index (values around one were considered acceptable).

Discriminant Analysis and Analysis of Variance

When a series of possible cluster solutions (e.g. 8. 7 or 6 groups) was identified, each of the solutions was then examined using step-wise discriminant analysis employing Wilk's Method (Nie et al, 1975). The predictor variables were initial factor scores and the dependent variables group membership. The overall prediction rates of each solution were compared to determine the best or optimal solution. Finally, for the ideal solution, group means and standard deviations were compared in order to determine the patterning or profiles among groups. In addition, a one-way analysis of variance was done on each derived factor by group. Scheffe's multiple comparison technique ($\alpha = 0.05$) was used to estimate real differences between groups on specific factors. Due to the methodology employed, any reported differences cannot be considered significant in the traditional sense. However, these tests provided estimates of the real distances between groups and provide insight into the distribution of these groups on specific factors on language ability.

Although cluster analysis capitalizes on error variability, by ascribing it to group membership when it is random, the overall efficiency of this solution may be judged by the prediction rates for group membership and the percentage of between-group variability. However, it should be stressed that these solutions are optimal in terms of variability and over-estimate between-group differences.

CHAPTER IV

RESULTS

Sample

Of the 72 medically diagnosed aphasics, forty were male. The mean age was 52.7 years with a standard deviation of 18.4 years and a range of 20 to 68. The average level of education was 10.2 years with a standard deviation of 3.6. Medical diagnosis was distributed as follows: 51 had cerebral vascular accidents; 19 had traumatic head-injuries, and 2 had brain damage related to anoxia. Time since onset (i.e. time from X to 0_1) ranged from one to fourteen months with the mean being 4.67 months. The mean treatment time was 5.1 months with a standard deviation of 2.4 months.

The obtained subtest means and standard deviations are reported in Table 5 for both pre and post-treatment. The means reported by Porch (1969) are also included in Table 5. Comparison of the pre-treatment means with Porch's standardized means revealed that the present sample differed significantly across the eighteen subtests (Hoteling $T^2 = 147.5$; $p < .001$). However, 95% simultaneous confidence intervals for each subtest spanned the corresponding mean in Porch's sample. Using a Spearman rank-order coefficient to assess the stability of the difficulty gradient across samples, a value of 0.94 was obtained. This result indicates that both samples encountered similar relative difficulty with the specific subtests.

For both subtests VIII and XI the sample means at 0_2 (see Fig. 2) exceeded 14.0; the variances were 1.45 and 0.31 respectively. Therefore,

TABLE 5

Means and Standard Deviations

Porch's Sample and Study Sample

Subtests[a]	Porch Xs	Pre-Treatment XPre	Pre-Treatment S.D.	Post-Treatment XPost	Post-Treatment S.D.	Difference Xs-XPre
I	8.40	7.91	3.42	9.88	3.89	+0.49
II	9.37	10.86	3.03	11.79	3.18	-1.49
III	10.36	11.89	2.89	12.67	2.92	-1.53
IV	9.45	9.20	4.11	10.65	4.08	+ .25
V	8.87	10.90	3.37	12.03	2.93	-1.03
VI	12.00	12.60	2.89	13.58	2.43	- .66
VII	9.94	11.72	3.09	12.85	2.42	-1.78
VIII	13.44	14.45	1.67	14.64	1.45	-1.01
IX	9.48	9.42	4.04	10.97	4.03	.06
X	12.67	13.27	2.56	13.81	2.01	- .60
XI	14.22	14.38	2.11	14.93	0.31	- .16
XII	11.17	11.83	3.97	12.90	3.26	- .63
A	6.01	6.64	2.75	7.95	3.02	- .63
B	6.62	7.99	3.60	9.88	3.54	-1.37
C	6.94	8.30	3.80	9.82	3.76	-1.36
D	7.39	8.91	3.93	10.18	3.83	-1.52
E	8.56	10.93	3.71	12.65	3.20	-1.37
F	10.49	11.99	3.31	13.36	2.46	-1.50
n	150	72		72		

[a]Porch's standard subtest labelling; descriptions provided in Chapter II, page 21.

these two subtests were exluded from subsequent analyses. For the remaining subtests, the means were well distributed over the scale; the variances were reasonably comparable among subtests and consistent over treatment (i.e. at O_1 and O_2). Therefore, sixteen subtests were retained for further analysis.

Factor Analysis

The intercorrelation matrix of these subtests is presented in Table 6. The first six eigenvalues obtained from the principal component analysis on post-treatment subtest scores were 10.50, 1.42, 1.26, 0.69, 0.63 and 0.30. The Kaiser-Guttman Rule suggested a minimum of three factors be retained while the Scree Test suggested a maximum of five factors be retained (Fig. 3). Therefore, unweighted least squares analyses were done on three, four and five factor solutions. The results of the principal component analysis and the three unweighted least squares analyses are presented in Table 7. The total percentage of accounted for variance implied by the principal component analysis (90.6%) and the dramatic increase in the number of residual correlations less than 0.15 in the five factor unweighted least squares analysis led to the choice of this solution for all subsequent analyses. Similarly, as this study was exploratory in nature, an overfactored solution was considered better than an underfactored solution. Therefore, the five factor solution was submitted to a series of Harris-Kaiser oblique/ orthogonal transformations (Hakstian and Bay, 1972).

The initial series of transformations (varying the degree of obliqueness (c) from 1.0 to 0.0 by 0.1 at each step) suggested that an

Table 6

Inter-Correlation Matrix of Post Treatment
P.I.C.A. Subtest Scores Used for Factor Analysis

Test	I	II	III	IV	V	VI	VII	IX	X	XII	A	B	C	D	E	F
I	1.00															
II	0.72	1.00														
III	0.69	0.89	1.00													
IV	0.92	0.68	0.64	1.00												
V	0.70	0.76	0.75	0.69	1.00											
VI	0.59	0.69	0.64	0.63	0.80	1.00										
VII	0.72	0.72	0.66	0.73	0.84	0.85	1.00									
IX	0.89	0.73	0.68	0.91	0.76	0.66	0.75	1.00								
X	0.56	0.66	0.60	0.63	0.68	0.83	0.85	0.66	1.00							
XII	0.66	0.55	0.54	0.72	0.43	0.47	0.52	0.73	0.53	1.00						
A	0.78	0.61	0.55	0.68	0.62	0.47	0.63	0.61	0.45	0.47	1.00					
B	0.75	0.60	0.53	0.74	0.66	0.57	0.71	0.64	0.53	0.49	0.83	1.00				
C	0.76	0.58	0.50	0.75	0.62	0.55	0.67	0.65	0.53	0.56	0.80	0.93	1.00			
D	0.77	0.67	0.59	0.76	0.73	0.63	0.73	0.70	0.59	0.62	0.80	0.88	0.91	1.00		
E	0.42	0.51	0.50	0.40	0.54	0.60	0.55	0.39	0.56	0.33	0.49	0.63	0.63	0.65	1.00	
F	0.29	0.43	0.44	0.24	0.46	0.44	0.40	0.26	0.41	0.28	0.44	0.46	0.46	0.52	0.77	1.00

FIGURE 3

Scree Test
(Principal Component Solution)

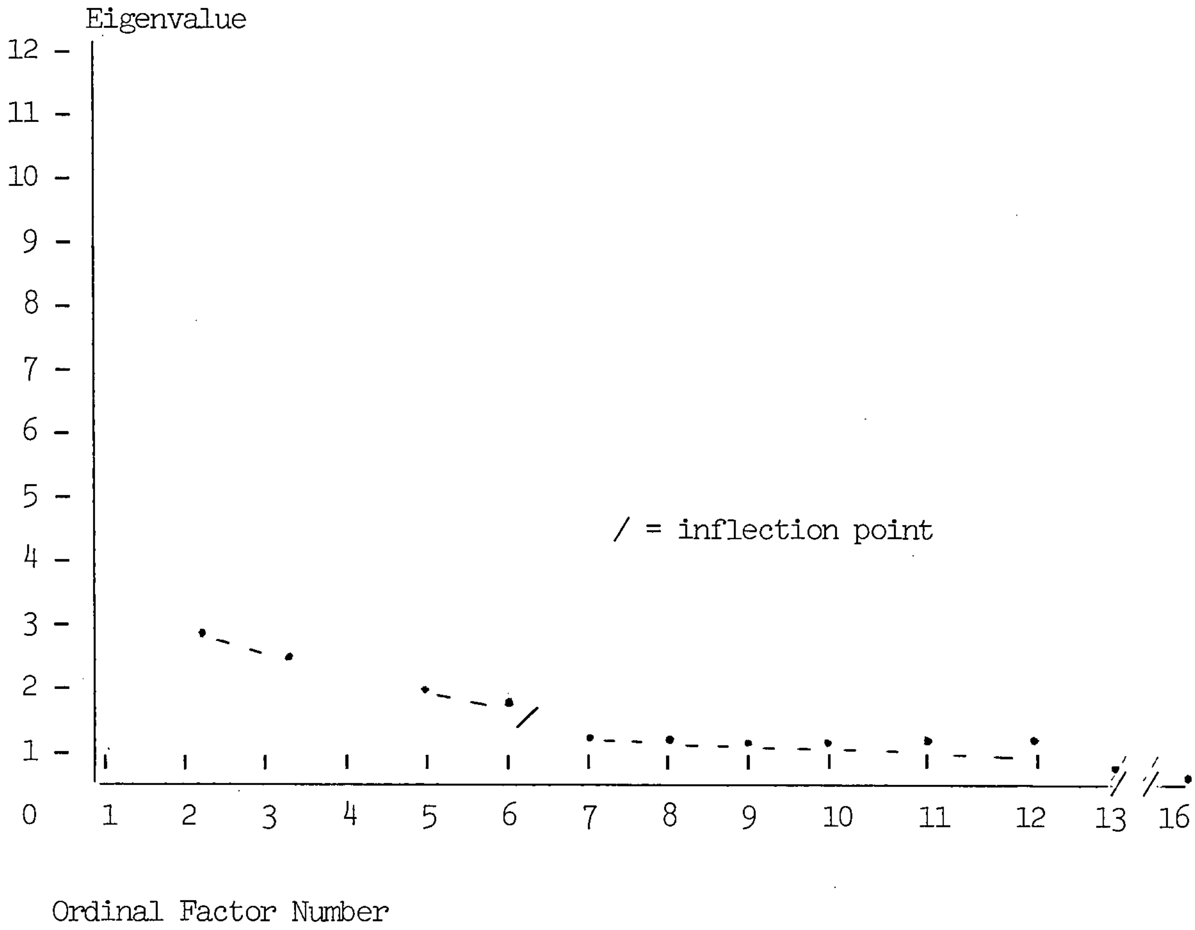


TABLE 7

Criteria for Estimating Number of Factors to be Retained

No. of Factors	Factor Method		
	Principal Component		Unweighted Least Squares
	Eigenvalues	Total Percentage of Accounted Variance	Percentage of Residual Correlations <0.15
1	10.50	65.6	a
2	1.42	74.5	a
3	1.26	82.4	52.3
4	0.69	86.7	54.7
*5	0.63	90.6	74.2*
6	0.30	92.5	a

*Solution of Choice

Footnote[a] - not done

TABLE 8

Results of Oblique Transformations

C	Hyperplane Count	Factorial Complexity
1.00	5	12
0.90	15	7
0.80	25	6
0.70	28	5
0.60	36	4
0.50	38	3
0.40	43	3
0.30	44	3
0.20	44	3
0.15	44	2
0.10	41	2
0.00	40	3

optimal solution fell between c = 0.10 and c = 0.20. Therefore, further transformations were done systematically varying c between 0.10 and 0.20. The optimal solution was found at c = 0.15. The hyperplane count and the number of factorially complex variables for each of the initial transformations and for c = 0.15 are presented in Table 8. At c = 0.15, the hyperplane count was at the maximum observed, 44, while the factorial complexity was limited to two.

The interfactor correlation and factor pattern matrices for this transformation are presented in Tables 9 and 10 respectively. The interfactor correlation matrix indicated that Factors 1, 2, 3, 4 are highly related and therefore were measuring in part a common dimension. Specifically, correlations among these four factors ranged from .60 to .70. However, Factor 5 appeared to be more independent in that the range of interfactor correlations between Factor 5 and each of the first four was 0.26 to 0.50. It appeared that Factor 5 was either a measure of a different dimension or language skill than the other four factors, or a measure of a non-language skill. To clarify this, the factor pattern matrix was inspected with special reference to both subtest task demands and Porch's hypothesized factor structure.

Factor 1 had high loadings on three of the four verbal subtests (IV, IX, XII); each of these variables had complexity one. The fourth verbal subtest, I, loaded on both Factor 1 and Factor 2. Therefore, the first factor was assumed to represent a measure of verbal fluency in that these four subtests require the subject to make a verbal response. Subtest I, with complexity 2, was the most difficult of the four verbal subtests.

Factor 2 had high loadings on four of the six subtests Porch labelled as graphic. Each of these subtests had complexity of one. These four subtests (A, B, C, D) all require the subject to write his responses. Therefore, this factor was assumed to reflect the patients' competency in expressing himself in written language.

Factor 3 had high loadings on three (VI, VII, X) of Porch's six gestural subtests. Each of these subtests had complexity of one. In addition, gestural subtest V loaded complexly on this factor and Factor 4. These subtests require a gestural demonstration of comprehension for either auditory or visually (reading) presented commands.

Factor 4 had high loadings on the remaining two gestural subtests (II and III). Each of these subtests had complexity of one. Both require the subject to demonstrate the function of presented stimuli and rely on the gestural modality of communicate the instructions. Although Factors 3 and 4 appeared to reflect somewhat different dimensions of communication skill, they had the highest inter-factor correlation (r = 0.70) and, in part, reflected Porch's gestural dimension may be further differntiated, dependent upon mode of instruction (i.e. verbal, reading or pantomime).

Factor 5 had high loadings on two remaining graphic subtests (E, F). Each of these subtests had complexity of one. Unlike the other four graphic subtests which require the patient to write, these two subtests require the patient only to copy letters to geometric shapes. Given the content of these two subtests and the relatively lower inter-factor correlations, it was hypothesized that these two subtests do not truly reflect language/communication abilities, but rather a visual/

TABLE 9

Inter-Factor Correlation Matrix

Factor	1	2	3	4	5
1	1.00				
2	0.69	1.00			
3	0.65	0.61	1.00		
4	0.67	0.60	0.70	1.00	
5	0.26	0.50	0.48	0.43	1.00

TABLE 10

Factor Pattern Matrix

Subtest	Modality	F-1	F-2	F-3	F-4	F-5
XII	Verbal	0.99	-0.16	-0.14	-0.08	0.25
IX	Verbal	0.70	0.05	0.19	0.15	-0.11
IV	Verbal	0.68	0.29	-0.13	0.02	-0.14
I	Verbal	0.48	0.45	-0.02	0.22	-0.17
A	Graphic	0.02	0.84	-0.09	0.02	-0.05
B	Graphic	0.05	0.95	0.08	-0.03	0.02
C	Graphic	0.17	0.86	0.02	0.16	0.12
D	Graphic	0.18	0.68	0.06	-0.00	0.16
VI	Gestural	0.04	-0.08	0.87	0.05	0.10
VII	Gestural	-0.01	0.22	0.82	0.04	0.10
X	Gestural	0.26	-0.20	0.82	0.08	0.15
V	Gestural	-0.14	0.26	0.52	0.42	0.09
III	Gestural	0.12	-0.11	0.06	0.93	0.10
II	Gestural	0.12	0.03	0.08	0.75	0.04
F	Graphic	-0.11	0.11	-0.05	0.15	0.75
E	Graphic	0.07	0.21	0.18	-0.02	0.72

perceptual process. Specifically, this factor correlated highest with factors requiring motor responses (rF2 = 0.50, rF3 = 0.48, rF4 = 0.43) and lowest with Factor 1 (verbal fluency, r = 0.26).

Hierarchical Analysis

Factor scores were derived on the pre-treatment scores using the procedure previously outlined and expressed in terms of T-score units ($\overline{X}$ = 50, S.D. = 10). These scores were then submitted to a hierarchical grouping analysis. For brevity, the last nine steps of this analysis are summarized in Table 11. The reason for only examining these last steps of the analysis becomes obvious when the specific aspects of variability are considered. First, at the ninth step only 14.5% of the total variability was partitioned into within-group variability. Second, the amount of variability added to within group variability from the tenth step to the ninth was only 1.2% of the total variability. These figures suggested that at this stage of the analysis, the existing groups were internally relatively homogeneous and that similar groups were still being combined.

As the analysis proceeded from nine groups to one, the increase of additional within group variability was marked, thus suggesting dissimilar groups were being combined. The percentage increase in within-group variability between steps is presented in the last column of Table 11. These figures suggested that until the six-group solution, relatively small portions of the total variability (i.e. 1.2 to 2.4%) were being partitioned into within group or error variability. After the six-group solution, this percentage rose more rapidly. Therefore,

TABLE 11

Results of Cluster Analysis

No. of Groups (Step No.)	Cumulative Error	Δ Error	Index	% of Total Variance	% of Change
9	5371.9	453.4	1.39	14.5	1.2
8	6095.2	723.3	5.36	16.5	2.0
7	6878.4	783.2	0.66	18.6	2.1
6	7769.2	890.8	0.96	21.0	2.4
5	9055.5	1286.3	2.66	24.5	3.5
4	11003.1	1947.7	2.57	29.8	5.3
3	14011.9	3008.8	2.18	37.9	8.1
2	20986.5	6974.6	3.95	56.8	18.9
1	36945.6	15959.1	2.57	100	43.2

the six-group solution was deemed best, in that after this step there was evidence that relatively homogeneous groups were not being combined. At this step only 21.0% of the total variability was within group variability; the remaining 79.0% was accounted for by between-group variability.

In order to gauge the overall group homogeneity for this solution and to compare its efficacy with solutions using fewer groups, a series of step-wise discriminant analysis using Wilk's Method was done on the 6, 5, and 4 group solutions using the derived factors scores as predictors. The respective prediction rates were 97.2%, 93.1% and 91.6%. This decrease in accuracy of prediction further suggested that the six-group solution still represented relatively homogeneous groups and that solutions using a smaller number of groups combined dissimilar groups. Therefore, the six-group solution was retained for further examination. The respective means, standard deviations of the factor scores and group sizes are presented in Table 12. (High scores indicated less impairment).

Except for Group VI (n = 4), the five remaining groups were of reasonable size (range 9 to 20) and therefore were assumed to be representative of real aphasic symptoms. Group VI was either a discrete and reasonably rare type of aphasic disorder or idiosyncratic to this sample. Therefore, with respect to the discussion of the P.I.C.A. and aphasic groupings, Group VI was considered separately.

The most striking feature of the five derived groups was that for each factor they appear to be distributed on a severity continuum and not a salient feature model. Specifically, for Factors 1, 3 and 4, the group scores, when rank-ordered according to group membership, were

TABLE 12

Means and Standard Deviations for
Five Factors on Six Groups

Group	n		Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
I	20	$\bar{X}$	59.1	59.5	57.1	56.8	55.0
		S.D.	3.18	4.13	2.02	3.35	3.30
II	9	$\bar{X}$	56.5	46.5	54.3	55.0	49.5
		S.D.	2.31	3.38	3.44	5.17	4.70
III	15	$\bar{X}$	45.0	46.9	52.8	48.3	56.5
		S.D.	4.92	7.03	3.75	4.66	3.90
IV	13	$\bar{X}$	39.4	39.8	41.5	43.7	46.9
		S.D.	3.32	4.10	8.48	8.03	5.97
V	11	$\bar{X}$	35.8	38.5	31.9	31.1	31.8
		S.D.	4.80	5.36	8.93	4.08	6.43
VI	4	$\bar{X}$	53.2	39.0	40.3	46.3	23.9
		S.D.	3.54	3.72	9.97	6.47	4.13

consistent across these factors for the respective groups. Factor 2 has only one 'cross-over; where Group III's score was slightly higher (+0.4) than Group II. In all probability, this was a function of sampling variability. Factor 5 also had one 'cross-over' but here Group III's score was the best of the five groups. The results, therefore, suggest that these groups represent an overall severity dimension of aphasic disorders and initial labels for the groups would be mild (i.e. Group 1) to severe impairment (i.e. Group V).

Univariate Analysis

In order to gauge the efficacy of the clustering solution on the specific factor dimensions, a one-way analysis of variance was done on each factor. The results of these analyses are sumarized in Table 13. As would be expected, highly significant between-group F-values were obtained. Similarly, the grouping procedure accounted for 86.0% of the total variability on Factor 1, 74.7% on Factor 2, 71.9% on Factor 3, 74.9% on Factor 4, and 83.1% on Factor 5. These figures, although inflated, highlight how much statistical power might be obtained by factoring subjects on impairment levels prior to experimental manipulation.

As aphasia theory has historically been based on univariate analysis procedures, the examination of groupings on specific factors was warranted. Using Scheffe's multiple comparison technique, the degree of difference between particular groups on specific factors was estimated. These comparisons are summarized in Table 14. On Factor 1, three distinct groups emerged. Groups I, II, and VI were similar and exhibit the least amount of impairment in verbal fluency. Groups IV and V were

TABLE 13

Analyses of Variance for the Five Factors

FACTOR 1					
Source	D.F.	Sum of Squares	Percent	F-value	F-prob.
Between Groups	5	5953.7	86.0	80.8	0.000
Within Groups	66	973.1			
Total	71	6926.7			
FACTOR 11					
Source	D.F.	Sum of Squares	Percent	F-value	F-prob.
Between Groups	5	4844.6	74.7	39.1	0.000
Within Groups	66	1637.2			
Total	71	6481.7			
FACTOR 111					
Source	D.F.	Sum of Squares	Percent	F-value	F-prob.
Between Groups	5	5967.9	71.9	33.8	0.000
Within Groups	66	2327.9			
Total	71	8295.8			
FACTOR IV					
Source	D.F.	Sum of Squares	Percent	F-value	F-prob.
Between Groups	5	5368.7	74.9	39.44	0.000
Within Groups	66	1796.9			
Total	71	7165.6			
FACTOR V					
Source	D.F.	Sum of Squares	Percent	F-value	F-prob.
Between Groups	5	7310.9	83.1	64.85	0.000
Within Groups	66	1488.2			
Total	71	8799.1			

also similar and had the severest amount of impairment in verbal fluency. Group III fell between these two clusters, being significantly more impaired than Groups I, II or VI, but significantly less impaired than Groups IV or V.

For Factor 2, the writing dimension, Group I exhibited the least impairment and was significantly better than the five remaining groups. Groups II and III appeared similar and were significantly less impaired than Group V. Groups IV and VI's performance was similar and fell between Groups II and III and Group V.

For Factor 3, the variable input modality gestural factor, the performance of Groups I, II, and III was similar and significantly less impaired than the other groups. Groups IV and VI's performance was similar, but Group IV performed significantly better than Group V.

For Factor 4, the pantomime - gestural dimension, the performance of Groups I and II was similar with Group I being significantly better than the other four groups. Group V was the most impaired group being significantly lower than all other groups. Groups VI and IV were similar and significantly more impaired than Group II. Group III fell between Group II and Group VI.

For Factor 5, the visual-perceptual factors, Groups I and III were least impaired and similar. In addition, these groups were significantly less impaired than Groups IV, V or VI. Groups V and VI were most impaired and significantly more impaired than all other groups. Group IV fell between these two extremes, being significantly less impaired than Groups V and VI, but significantly more impaired than Groups

TABLE 14

Scheffe Multiple Comparisons on Factors by Group
(Homogenous Groups are Underlined)

	Group[a]					
Factor	High					Low
I	I	II	VI	III	IV	V
2	I	III	II	IV	VI	V
3	I	II	III	IV	VI	V
4	I	II	III	IV	VI	V
5	III	I	II	IV	V	VI

I and III. Group II was similar to Groups I and IV, but significantly more impaired than Group III.

Qualitative Statements Concerning Groupings

Group 1: This group was the least impaired group on four of the five factors. On Factor 5, Group II performed better, but not significantly better. Therefore, Group I may be termed as showing mild impairment on all five dimensions of language ability.

Group II: Group II exhibited significantly more impairment than Group I on only the second factor. On all other factors the mean scores were lower, but not significantly. These findings would suggest that this group may be described as exhibiting mild to moderate impairment of language abilities.

Group III: This group exhibited significantly more impairment than Group I on three of the five factors, Factors 1, 2 and 4. On three of the five factors, Groups II and III were similar. On Factor 1, Group III showed significantly more impairment but on Factor 5, this trend was reversed. However, as the relationship between Factor 5 and language abilities per se was tenuous, this group may be labelled as exhibiting moderate impairment.

Group IV: This group exhibited significantly more impairment on all five factors than Group I. More importantly it performed at a significantly lower level than Group III on three of the five factors. These findings would indicate that this group exhibited moderate to severe impairment.

Group V: On all five factors, this group exhibited significantly more impairment than Groups I, II or III. On Factors 3, 4 and 5, this group's performance was significantly more impaired than Group IV. Excluding Group VI, this group had the lowest mean score on all five factors. Therefore, this group may be described as exhibiting severe impairment of language abilities.

Group VI: Given the small size of this group (n = 4), the results on this group must be viewed cautiously as they may be idosyncratic to this particular sample. However, this group had the only profile which was consistent with a salient feature or deficit model of aphasia. Specifically, this group was similar to Groups I and II on the verbal fluency factor, thus suggesting mild to moderate deficits in this area. However, on Factor 5, the mean score of 23.9 was suggestive of gross impairment and was similar to the mean obtained by Group V, only lower. On Factors 2, 3 and 4, this group's mean score fell between Groups IV and V and thereby also performed at significantly lower level than Group I. Therefore, this group may be described as exhibiting mild impairment in verbal fluency, moderate to severe deficits in writing and gestural skills and gross impairment in copying tasks.

CHAPTER V

DISCUSSION AND CONCLUSION

Introduction

The purpose of this study was to explore a methodological procedure for the examination and identification of sub-populations with an atypical population. Therefore, this discussion is divided into two parts. The first considers the findings of this study within the context of current aphasia theory and, in particular, instrumentation. The second attempts to illustrate the place and potential of this type of analysis procedure in the initial assessment of other atypical populations. The results of this study are used as concrete examples of the research potential of this type of analysis.

Limitations of this Study

Before examining the actual findings and implications of this study to aphasia theory and research, certain limitations within this study should be reviewed. First, the limited sample size suggests that the findings should be viewed cautiously in terms of generalization. Accordingly, experimental replication across different samples is advised before using the outlined categorization system and derived factor domains in a treatment-oriented research design. Second, the categorization system only pertains to the P.I.C.A. and therefore may be inappropriate for other aphasia assessment instruments. Third, the statistical procedures used are designed to give optimal or maximal solutions within the constraints of the data. Therefore, the reported results are, in all probability, estimates of upper, not lower, bounds.

With respect to this sample representing an aphasic group, the subtest means suggest language impairment consistent with the diagnosis of aphasia. In addition, when compared to Porch's standardization sample, this group did not score significantly differently on any specific subtest although over all comparisons a significant Hotelling's T^2 was obtained. However, two subtests (VIII and XI) were discarded because of observed ceiling effects. Therefore, only sixteen of Porch's eighteen subtests were included for further analysis.

Factor Analysis

The factor analysis procedure employed suggested that a five-factor solution would be best for the data collected in this study. This solution accounted for 90.6% of the total variance among the sixteen post-treatment subtests, thus indicating that the skills assessed by this battery may be described by substantially fewer than sixteen dimensions or tests. The exclusion of subtests VIII and XI split Porch's gestural dimension into two highly correlated factors. In the factor analysis of Porch's original data, (Clerk et al, in press), these two subtests (VIII and XI) anchored Porch's gestural dimension (their respective loadings being 0.96 and 0.89). This 'anchoring' is statistically related to the original lack of variance in these two subtests. In the present study, the exclusion of these two subtests yielded two gestural factors, which in turn led to a reduction of factorial complexity among the gestural variables. Specifically, in the present analysis only one of the six gestural subtests was complex; while the factor analysis of Porch's standardization sample yielded four factorially complex gestural variables.

This five factor solution also separated Porch's graphic dimension into two factors. The four graphic subtests which required a written response (A, B, C, D) defined one dimension, while the two graphic subtests (E, F) requiring a copying response, identified the second. Other than the differentiation of the two hypothesized factors into four sub-factors, the results of this factor analysis were remarkably consistent with both Porch's formulation and the factor analysis performed on Porch's initial data for his test battery. Specifically, with respect to Porch's formulations, a verbal, two gestural and two graphic dimensions could be identified. More importantly, only one subtest (I) loaded across modalities (i.e. having relatively high loadings on both verbal and gestural dimensions). Although it was factorially complex, subtest V loaded on the two gestural dimensions.

The five derived factors are consistent with previously suggested dimensions of aphasic disorders. Factor 1, the verbal fluency dimension, is clearly related to Howes fluent/dysfluent continuum or dichotomy (Howes, 1964). Factor 2, the writing dimension,has been clinically labelled agraphia, the inability to write words (Walton, 1974). More importantly, disruption in this skill has been associated with three discrete types of aphasia, namely, transcortical sensory aphasia, transcortical motor aphasia and alexia with agraphia (Goodglass and Geschwind, 1975). Factor 3, a gestural output modality using a variety of input modalities, supports the multi-modality model of Wepman and Jones (1961). Factor 4, the pure gestural dimension, is in accord with the assessment of pantomime in aphasia (Goodglass, 1963). Factor

5, the copying dimension, may not represent a true dimension of aphasia. Specifically, the inability to copy designs or letters does not require comprehension or expression with a symbolic language model. However, the dimension may represent a general apraxia, the inability to carry out a 'willed' motor response even though motor and sensory pathways are intact (Walton, 1974).

The findings that (1) the derived factors can be related to behavioural deficits associated with aphasia; (2) this and the previous factor analysis done on this battery are relatively consistent; and (3) the support for Porch's hypothesized dimension support the construct validity of the P.I.C.A. However, what is of further interest in regard to this battery is the magnitude of the interfactor correlations, both in this and the previous study. For this sample, the factors were more highly correlated than for the previous one. Statistically, these higher interfactor correlations are in part accounted for by the increase in factors (5 versus 3) which reduced the factorial complexity and allowed specific dimensions to be better defined. For example, the highest inter-factor correlation was between the two gestural factors ($r = 0.70$). Therefore, although a higher order factor analysis was not done, the high intercorrelations between Factors 1 to 4 again suggested a higher order factor of general language abilities.

Factor 5 appeared to be somewhat of an outrider to this model. This factor has a relatively low correlation with Factor 1 ($r = 0.26$) and, although the magnitude of the remaining inter-factor correlations with Factor 5 were higher (0.50, 0.48, 0.43), they were still lower than all other inter-factor correlations (range 0.60 to 0.70). Perhaps Factor 5 requires a non-language skill (e.g. visual acuity) together with a motor response congruent with Factors 2, 3, and 4. Such a hypothesis would call into question the validity of subtests E and F as measures of language abilities.

Hierarchical Analysis

The hierarchical grouping analysis added further support for a general severity model of aphasia. Altogether six groups were identified. Five of these were ordered upon a severity continuum, not a salient features model. For these five groups significant changes in ordinal ranking of means only occurred once on the five factors. In addition, this 'significant' crossover occurred on Factor 5.

The sixth identified group was inconsistent with this severity model. However, this group was small (n = 4) and its lowest mean was on Factor 5. In addition, this group was equivalent to the least impaired groups on the verbal fluency factor and exhibited moderate impairment on the remaining factors. Therefore, the question arises whether this group's most salient deficit was 'aphasia'. In other words, individuals in this group may have exhibited mild aphasic symptoms, but in addition they may have had more extensive damage to other brain mechanisms. Given the low representation within this group, any statements concerning it must be viewed judiciously.

Besides the qualitative aspects of this grouping procedure, overall efficacy in terms of variability considerations and prediction rates were extremely impressive. In terms of overall variability, the six-group solution suggested that the groups were relatively homogenous, with 79.0% of the total variability between groups. The multiple discriminant analysis prediction rate of 97.2% provided further confirmation, suggesting there was virtually no overlap between groups in multi-variate space.

Grouping and Current Aphasia Theory

When specific factor dimensions relative to derived groups are examined, the relationship between these dimensions and current aphasia models becomes apparent. As mentioned earlier, Factor 1 clearly represents Howe's verbal fluency dimension. More importantly, if one examines the relative group sizes and their distribution along the factor, it is clear how a dichotomous model of verbal fluency/dysfluency can be evolved. In addition, given the reported F-value and the number of significant differences found, a two-group analysis (i.e. fluent/dysfluent) would in all probability produce a significant difference. Specifically, Groups I, II and VI ($n = 33$) would be Howes 'fluent' aphasics while Groups IV and V ($n = 24$) would be termed 'dysfluent'. Group III ($n = 15$) does fall between these two categories, but in Howes' model the high scores would be fluent, the low scores dysfluent. However, by using the grouping procedure employed here, it is apparent that fluency skills in fact are distributed along a continuum and are not purely dichotomous.

Similar arguments can be made for the writing and the two gestural dimensions in terms of their representation in aphasia models. Accordingly, the findings indicated the limitations of examining aphasia on a uni-variate model with respect to developing a theory of aphasia.

Conclusions Relevant to the P.I.C.A.

Although the P.I.C.A. is a well-standardized and recognized aphasia battery, the results obtained in this study suggested that at most it measured five dimensions of language/communication ability and that these dimensions were highly inter-related. Such findings may indicate that the P.I.C.A. is a more elaborate test than necessary, in that many of the subtests may be redundant or provide minimal unique information with respect to communication abilities per se. Specifically, subtests VIII and XI appeared to be too easy for both the subjects in this sample and in Porch's sample. Subtests E and F did not appear to be directly related to symbolic language formulation. With the exclusion of these four tests, the battery still uses fourteen subtests to assess a suggested continuum of impairment.

For the clinician these results suggest that particular attention be paid to the level of difficulty when monitoring treatment gains. Specifically, the noted ceiling effect on subtests VIII and XI may 'wash out' improvement in gestural skills. In effect, using these two subtests is similar to adding a constant, and thus they may inflate initial mean scores. For example, if an individual scores perfectly on all eight gestural tasks, the overall gestural mean is 15.0. In comparison, another individual may score ten on the six other gestural tasks and fifteen on subtests VIII and XI (not an unlikely set of scores) and thus obtain an overall mean of 11.25. Because of the ceiling on tests VIII and XI, the real difference between these two individuals is lessened (3.75 units not 5.0). More importantly, if the second individual improves on each of the six subtests by 2.0 units, the overall mean can only rise by 1.5 units, thus under-estimating the real improvement. As the subtests appear to be highly reliable, it may be worthwhile to consider them independently in terms of retest. Or, after the initial assessment the clinician could select a group of subtests which would be repeated on retest based on a subject's performance. Although this procedure would violate standardization procedures, it would provide a more accurate estimate of true gains and be more economical in terms of time. Two concerns of this procedure are statistical regression and maturation or, specifically, spontaneous recovery. However, changes measured by the instrument which are a result of spontaneous recovery, do not indicate that the instrument is not measuring real changes rather causation is undermined. However, this procedure is susceptible to statistical regression towards the mean and the clinician therefore should select both above and below the standardization mean whenever possible.

To the researcher, these results suggest that initial grouping of an aphasic sample should be done on severity continuum if the P.I.C.A. is the dependent measure. In addition, the factor analysis suggested at a maximum the P.I.C.A. measures five dimensions of communication abilities not eighteen. This reduction in dependent variables gains considerable statistical power and reduces the potential overall experiment-wise error rate. In addition, model or conceptual development is less cumbersome. Initial subests means and standard deviation should be reviewed with respect to ceiling effects as specific subtests appear too easy for all subjects.

METHODOLOGICAL CONCERNS

Introduction

With respect to atypical populations in general, the outlined procedures may provide insight into better delineation of these populations. The aphasic sample used here for explanatory purposes has much in common with other atypical populations. Specifically, aphasics can be differentiated from the general population due to deficits in language abilities. Although these deficits are related to brain-behaviour relationships, many other populations exist which can be differentiated from the general population based on specific areas of deficits. Examples of these populations are the deaf, reading disabled, learning disabled, minimally brain damaged, epileptics or psychiatric populations. For each of these populations, it is possible to differentiate the population from normals and usually there is one main dimension for this discrimination. Once the population is identified, however, it violates the 'true' experimental paradigm in the same manner as the aphasic population. Therefore, the question of intra-population distribution becomes important with respect to research into remediation. Dependent upon the historical level of enquiry into such populations, the researcher must chose the appropriate level for initial investigation. For example, aphasiologists have developed several conceptual models of classification as well as some psychometrically appealing test batteries. By using multivariate techniques it was possible in this study to examine these conceptual classification models and an empirically derived classification model to be explored and further developed.

Factor Analysis

The power of multivariate designs is related to the number of dependent variables, groups and subjects. Ideally, the number of dependent variables should be minimized and yet representative of the domain of skills being sampled (e.g. language/communication abilities). Factor analysis allows the researcher to fullfil these restraints, given that the initial test battery is reliable and at some point valid. Ideally, a set of independent factor score coefficients enhances the acceptability of such a procedure. In the case reported here, the factor score coefficients are not independent but outside validity measures indicate that the factor structure of the P.I.C.A. is reasonably stable and consistent with hypothesized dimensions.

One issue which has long plagued factor analysts is how to determine the number of factors to be retained. This study established a minimum by using the Kaiser Guttman Rule and maximum by using a Scree Test from the right side of the eigenvalue plot. Possible solutions were then submitted to an unweighted least squares analysis and the residual correlation matrices examined. This procedure suggested that five factors, two more than suggested by the Kaiser-Guttman Rule, best represented the dimensions within the P.I.C.A. This study should emphasize the necessity of using multiple criteria when determining the number of factors to be retained. In addition, a further criteria of interpretability should be applied to the final factor pattern matrix. Overall, the derived factors and the variable loadings were conceptually consistent in this study.

Hierarchical Analysis

Once the minimum number of representative dependent variables has been delineated, the second means of gaining statistical power in a multivariate design is by minimizing the number of groups. However, because of variance and covariance considerations, these groups should be as internally homogeneous as possible. In this particular case, the cluster analysis suggested that a six-group solution best fit these criteria. Specifically, the number of potential groups was reduced from 72 to 6, with only 21% of the total multivariate variability within the groups. More importantly, this solution indicated that subjects should be grouped on a severity continuum rather than in terms of a salient features model. This finding in itself is of benefit in that the salient features model is the accepted grouping procedure, both for research and clinical practice. Therefore, in addition to providing an 'optimal' grouping procedure, this analysis also raised questions concerning the overall conceptual validity of many aphasia models.

Although cluster analysis does capitalize upon error, thus inflating the between-group variability estimates, this aspect of the analysis does not affect conclusions drawn from the group profiles. In this particular study, the step-wise discriminant analysis suggested the amount of overlap in multivariate space between these six groups was minimal. Again, however, experimental replication should be stressed as the means for establishing both the stability and generalizability of these results.

Directions for Future Research

As with all experimental studies, the first recommendation for future research is replication of this study. As the findings in this study are strong in terms of aphasia, replication would provide estimates of their stability and generalizability. If the findings were found to be both stable and appropriate for other aphasic samples, then a model of grouping for treatment could be evolved to evaluate differential treatment efficacies.

The second limitation of these findings is generalizing from these results to other aphasia instruments. Although the P.I.C.A. is psychometrically the best instrument to evaluate aphasia, it may be conceptually biased towards a severity model of aphasia. Therefore, to develop a statistical or empirical model of aphasia, other aphasia instruments should be analyzed using these procedures. If the groupings or group profiles appear to be a function of instrumentation, then one must conclude that the instruments are conceptually biased. If, however, the 'groupings' appear to be a function of severity or impairment, then one must conclude that aphasia is a general deficit affecting language abilities. If the latter conclusion is found, then there are definite implications for the differentiation of clinical/treatment and brain/behaviour studies or research.

In addition, the patients used in this study had traumatic insult to the brain. The nature of the insult suggests that large areas of the cortex were probably involved and therefore one may assume that the behavioural losses were accordingly more general than in the case of a discrete small lesion (e.g. epilepsy). However, traumatic histories now account for the vast majority of aphasics.

Finally, this study raises some question concerning the number of useful subtests on the P.I.C.A. and the number of factors or dimensions of language dimensions being measured by this instrument. By excluding two 'ceiling' gestural subtests, the initial postulated gestural dimension subdivided into two factor domains. With this reduction in subtests the graphic dimension was also subdivided as subtests E and F had more weight in the analysis (12.5% rather than 11.1% of the total variance). This breaking of the graphic dimension into two makes more inherent conceptual sense as subtests E and F do not truly require language abilities. Moreover, these results suggest that subtests VIII and XI add minimal information to clinical interpretation of the P.I.C.A. and may in fact distort real differences among subjects and hide improvement for a specific subject. Therefore, these results suggested that a revision of the P.I.C.A. could provide a more economical instrument in terms of time, and a more accurate instrument in terms of assessing the degree of language impairment and the degree of improvement following treatment.

APPENDIX A

Factor Analysis of Porch's Standardization Sample

APPENDIX A

TABLE 1

Factor Pattern Matrix
Harris Kaiser Oblique Transformation

c = 0.3

Group	Test	Factor 1	Factor 2	Factor 3
Verbal	1	0.78	0.21	0.05
	IV	0.88	0.13	0.06
	IX	0.87	0.13	0.06
	XII	0.76	0.04	0.18
Graphic	A	0.13	0.83	0.00
	B	0.09	0.92	0.00
	C	0.00	0.95	0.04
	D	0.12	0.86	0.06
	E	0.01	0.68	0.28
	F	0.07	0.37	0.59
Gestural	II	0.15	0.25	0.62
	III	0.17	0.24	0.60
	V	0.12	0.44	0.40
	VI	0.36	0.02	0.65
	VII	0.11	0.39	0.49
	VIII	0.13	0.05	0.96
	X	0.37	0.08	0.67
	XI	0.09	0.16	0.89

APPENDIX A

TABLE 2

Factor Correlation Matrix

Factor	1	2	3
1	1.00		
2	0.51	1.00	
3	0.52	0.50	1.00

REFERENCES

Carroll, J.A. A factor analysis of verbal abilities. Psychometrika, 1941, 6, 279-307.

Cattel, R.B. The scree test for the number of factors. Multivariate Behavioural Research, 1966, 1, 245-257.

Clark, C., Crockett, D., & Klonoff, H. Factor analysis of the Porch Index of Communication Ability. Brain and Language (in press).

Crockett, D. Multivariate comparison of Howes' and Weisenburg and McBride's models of aphasia on the Neurosensory Center Comprehensive Examination. Perceptual and Motor Skills, 1976, 43, 795-806.

Eisenson, J. Examining for aphasia. New York: Psychological Corporation, 1954.

Eisenson, J. Aphasia in adults: basic considerations. In Travis (Ed), the Handbook of Speech Pathology and Audiology, New York, Appleton-Century-Crofts, 1971.

Fisher, R.A. The design of experiments. London: Oliver and Boyd, 1935.

Gerstein, J.S. Rehabilitation potential. In J. Licht (Ed), Stroke and Its Rehabilitation, Baltimore: Waverly Press Inc., 1975.

Golden, J. Dynamics of head injury. Journal of Physical Rehabilitation, 1977, 15, 59-74.

Goodglass, H., & Geschwind, N. Language disorder. In Jones (Ed) the Handbook of Perception, Speech and Language, New York: Academic Press, 1975.

Goodglass, H., Hyde, M., & Blumstein, S. Frequency picturability and the availability of nouns in aphasia. Cortex, 1969, 3, 104-112.

Goodglass, H., & Kaplan, E. The assessment of aphasia and related disorders. Philadelphia: Lea and Febiger, 1972.

Gorsuch, R. Factor analysis. Philadelphia: Saunders, 1974.

Guttman, L. Some necessary conditions for common factor analysis. Psychometrika, 1954, 19, 149-184.

Hakstian, A.R. A computer program for oblique factor transformation using the generalised Harris-Kaiser procedure. Educational and Psychological Measurement, 1970, 30, 703-736.

Hakstian, A.R., & Bay, K.S. Alberta general factor analysis programme (AGFAP), University of Alberta: Division of Educational Research Services, 1972.

Halstead, W.C., & Wepman, J.M. The Halstead-Wepman aphasic screening test. Journal of Speech and Hearing Disorders, 1959, 14, 9-15.

Head, J. Aphasia and kindred disorders of speech, Vol. 1 & 11, London: Hafner, 1926.

Howes, D. Application of the word-frequency concept to aphasia. In (A.V.S. de Reuch and M O'Connor (Eds), IBA Foundation Symposium on Disorders of Language, Boston: Little Brown and Company, 1964.

Howes, D., & Geschwind, M. Quantitative studies of aphasic language. In D. Rioch and E. Weinstein (Eds), Disorders of Communication, Baltimore: Williams and Wilkins, 1964.

Klonoff, H., Low, M., & Clark, C. Head injuries in children: a prospective five year study. Journal of Neurology, Neurosurgery and Psychiatry, 1977, 40, 1211-1219.

Lezak, M. Neuropsychological assessment. New York: Oxford University Press, 1976.

Luria, A. Higher cortical functions in man. New York: Basic Books, 1966.

Milner, B. Interhemispheric differences in the localization of psychological proccesses in man. British Medical Bulletin, 1972. 17, 272-77.

Morrison, D. Multivariate statistical methods. New York: McGraw Hill, 1975.

Naeser, M.A., & Hayward, R.W. Correlations beteen size and density of CT scan lesion sites and severity of aphasia in stroke patients. In Abstracts: Academy of Aphasia 15th Annual Meeting, Montreal, 1977.

Nie, N., Hull, C., Jenkins, J., Steinbrenner, K., & Bert, D. Statistical package for social sciences (2nd Ed). New York: McGraw Hill, 1975.

Osgood, C.E., & Miron, M. (Eds). Approaches to the study of aphasia. Urbana: University of Illinois Press, 1963.

Patterson, J., & Whittaker, J. Hierarchical grouping analysis with optional contingency constrain. (U.B.C: GROUP), Vancouver: U.B.C. Computer Centre, 1973.

Penfield, W., & Roberts, L. Speech and brain mechanisms. Princeton: Princeton University Press, 1966.

Porch, B. Porch index of communication ability: Theory and development (Vol. 1), Administration scoring and interpretation (Vol. 11). Los Angeles: Consulting Psychologists Press, 1968.

Schuell, H., Jenkins, J., & Carroll, J. A factor analysis of the Minnesota test for differential diagnosis of aphasia. Journal of Speech and Hearing Research, 1962, 5, 349-69.

Scheull, H., Jenkins, S., & Jiminez Pabon, E. Aphasia in adults. New York: Harper & Row, 1969.

Walker, A.E., Caveness, W.F., & Critchley, J.D. The late effects of head injury. Springfield: C.C. Thomas, 1969.

Walton, J.N. Essentials of neurology. Toronto: J.B. Leppincott Co., 1974.

Weisenburg, T., & McBride, K.E. Aphasia. New York: The Commonwealth Fund, 1935.

Wepman, J., & Jones, J. Studies in aphasia: an approach to testing. Chicago: Education Industry Service, 1961.

Whittaker, H. Electrical stimulation mapping of language in the cerebral-cortex. In Abstracts: Academy of Aphasia 15th Annual Meeting, Montreal, 1977.

Williams, J.M. Synesthetic adjectives: a possible law of semantic change. Language, 1976, 52, 461-478.

Winner, E., & Gardner, H. The comprehension of metaphor in brain damaged patients. Brain, 1977, 10, 717-729.

Zurif, E., Caramazzo, A., & Galvin, G. Semantic feature representations for normal and aphasic language. Brain and Language, 1974, 7, 167-187.

Zurif, E., Caramazza, A., & Myerson, R. Grammatical judgments of agrammatic aphasics. Neuropsychologia, 1972, 10, 403-417.

KLONOFF, H., CLARK, C. "Measurement of Staff Attitudes Towards Computerization", Hospital and Community Psychiatry, 26: 823-825 (1975).

KLONOFF, H., CLARK, C. "Drug Patterns in the Chronic Marijuana User". The International Journal of the Addictions, 11: 71-80 (1976).

KLONOFF, H., CLARK, C., HORGAN, J., KRAMER, P. and McDOUGALL, G. "The MMPI Profile of Prisoners of War", Journal of Clinical Psychology, 32: 623-627 (1976).

CROCKETT, D.J., KLONOFF, H., CLARK, C. "The Effects of Marijuana on Verbalization and Thought Processes", Journal of Personality Assessment, 40: 582-587 (1976).

KLONOFF, H., McDOUGALL, G., CLARK, C., HORGAN, J. and KRAMER, P. "The Neuropsychological, Psychiatric and Physical Effects of Prolonged and Severe Stress - Thirty Years Later", Journal of Nervous and Mental Disease, 163: 246-252 (1976).

KLONOFF, H., LOW, M.D. and CLARK, C. "Head Injuries in Children: A Prospective, Five Year Follow-up", Journal of Neurology, Neurosurgery and Psychiatry, 40: 1211-1219 (1977).

CLARK, C., CROCKETT, D.J. and KLONOFF, H. "Factor Analysis of the Porch Index of Communication Ability", Journal of Brain and Language. April 1977.

CLARK, C., CROCKETT, D.J. and KLONOFF, H. "A Multivariate Model for Analyzing the Effects of Treatment in Aphasic Patients". (Submitted for publication).

KLONOFF, H., CLARK, C. and TAGGART, M. "A Hospital Computer System". (Submitted for publication)