ARTICLE

Behavioral Data and Early Evaluation of Treatment Outcome

Barbara Ann Johnson
National Technical Institute for the Deaf, Rochester, New York

Abstract

Clinicians who treat aphasic people are mandated to design and implement treatment programs that maximize the benefits of intervention as expeditiously as possible to assure treatment cost effectiveness, a process which could be done more predictably if clinicians had access to methods for evaluating the outcome of treatment tasks based on the initial collection of behavioral data. The quarter-intersect method of estimating trend and predicting performance has proven itself to be a quick and easy method for estimating the trend of serial data. Therefore it was applied to consecutive performance scores from 16 treatment tasks that comprised the therapy programs for five aphasic gentlemen. The primary findings of the study suggest that the quarter-intersect method is not only quick and easy, but also yields valid results comparable to those derived from the complicated and time-consuming least-squares statistical procedure. Further analysis of the data showed that, using the quarter-intersect method, 6 and 10 sessions of treatment accurately predicted whether progress would occur in 28 treatment sessions 75% of the time, and that 14 sessions of treatment predicted whether progress would occur 93% of the time. The implications of these findings are discussed with regards to making timely patient management decisions.

Introduction

For several years questions surrounding the efficacy of aphasia treatment have been raised and debated (Benson, 1979; Darley, 1972, 1982; "Prognosis for Aphasia," 1977; Sarno, 1970). Initially the argument concerned whether or not treatment offered measurable gains in communication skill. Although no single study provided conclusive evidence to support behavioral treatment of aphasia based on treatment gains, in 1982 an accumulation of studies were cited, and collectively they provided substantial reassurance that aphasia treatment is beneficial to most aphasic patients who participate (Darley, 1982). Yet, the controversy continues (Wertz, 1986).

Individuals with aphasia differ in their characteristic symptoms, potential for progress, and ultimately in their rate and pattern of change during recovery. In addition to these individual patient differences, a variety of treatment approaches and task types exist, and these also have potential to influence the outcome of a person's treatment program. For example, not all tasks or treatment strategies work; other strategies appear to work well with some patients but not with others, or appear to be more effective at some times than at others. Further, few data-based suggestions are available to clinicians who are interested in evaluating the value of ongoing treatment programs on an individual basis. For example, if a clinician is interested in evaluating the potential effect of a recently initiated treatment program, he can not give a data-based response to many questions that pertain to anticipated communication gains. One question may be: How many times does the patient need to participate in the task before we are reasonably sure that he will make progress in it and thus improve his communication skill?

Even without satisfactory answers to this question and many similar questions, clinicians who treat aphasic people are mandated to design and implement treatment programs that maximize the benefits of individualized intervention as expeditiously as possible to assure the cost-effectiveness of aphasia treatment. Perhaps this could be done more effectively if clinicians had access to methods for evaluating the eventual outcome of treatment tasks based on the initial collection of behavioral data.

The purpose of this study is to assess the validity of one specific data-based method for evaluating series of behavioral data, and to explore its utility in evaluating the eventual outcome of treatment tasks based on the initial collection of behavioral data. Some suggestions for using the method in evaluating individualized aphasia-treatment programs, and specific criteria for continuing, terminating, or altering a treatment task, are presented and discussed.

Method

Subjects

Five male veterans, all right-handed, were diagnosed as aphasic by speech and language testing and by the judgment of a certified speech-language pathologist. The *Porch Index of Communicative Abilities* (Porch, 1967) and the picture-presentation task of the *Boston Diagnostic Aphasia Examination* (Goodglass & Kaplan, 1972) were included in the initial evaluation in order to classify subjects according to the severity and fluency of their aphasia. In Table 1, subjects are described according to these and other relevant variables, including onset time and etiology.

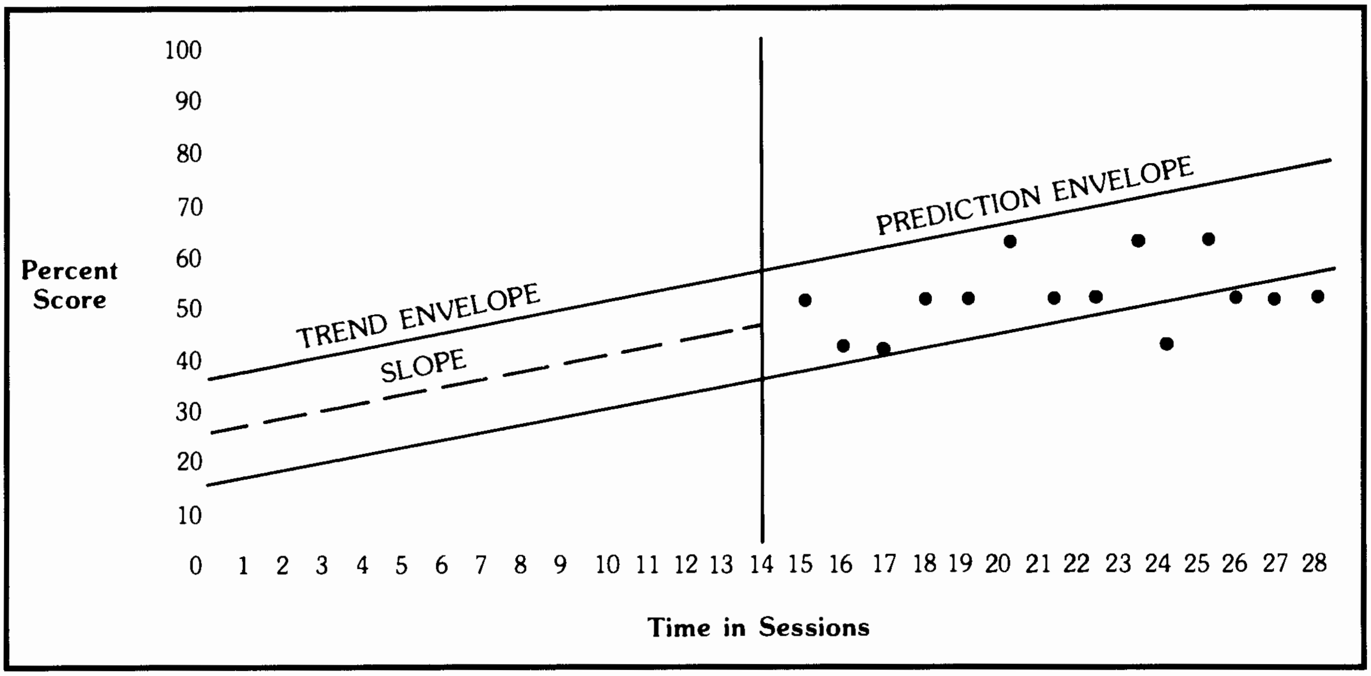
Data Collection

Individualized treatment programs were developed for each subject; scoring systems were designed for all treatment tasks; and 2 to 14 baseline measurements were made for each task to establish initial performance levels and eliminate tasks with ascending baseline scores. Two to four tasks were selected for each subject's treatment program, and a total of 16 treatment tasks were included in the data analysis. Treatment was scheduled for two to four one-hour sessions per week, depending on

Table 1. Patient Characteristics

Variable	F.P.	F.B.	R.E.	D.M.	H.E.
Fluency	Nonfluent	Fluent	Fluent	Nonfluent	Fluent
Severity	Moderately impaired	Mildly impaired	Mildly impaired	Severely impaired	Functional
Description of Aphasia	Severe verbal formulation difficulty & severe phonological selection & sequencing difficulties, occasional verbal paraphasia, neologistic jargon, mild auditory comprehension deficit	Moderate phonological selection & sequencing impairment	Mild fluent aphasia, mild auditory comprehension deficit, word finding difficulties, apraxia, conduction aphasia	Severe receptive deficit, severe to profound expressive difficulties, severe apraxia	Mild to moderate auditory & visual comprehension & word-finding difficulties
Associated Defects	Hemiparesis of the right leg, paralysis of right arm	Right-sided weakness	Resolved right hemiparesis	Right hemiplegia	Right hemiparesis
Age at Onset	56 years	52 years	48 years	45 years	69 years
Premorbid Education	Business School	Grade 9	Grade 12	Grade 12	Dental School
Premorbid Intelligence	Normal	Normal	Normal	Normal	Superior
Premorbid Language Function	Normal	Normal	Bilingual English/Spanish	Normal	Superior
Occupation	Business administrator in construction	Beer distributor	Power plant project officer (Navy)	Machinist	Dentist
Social Milieu	Married, 5 grown children	Married, 2 teenagers at home	Married, 8 year old son at home	Married, 5 children D.M. in convalescent home	Bachelor, home for the elderly
Medical History	Hypertension, cardiac history, rheumatic arthritis	Hypertension, cardiac failure, TIA	Not remarkable	Hypertension	Not remarkable
Associated Illness During Recovery	Compactions	None reported	None reported	Polycystic kidney disease (required dialysis)	Diabetes
Cause of Aphasia	Left CVA, 5/80	Left MCA thrombotic CVA, 4/79	Left CVA, 11/79	Left CVA, 11/80, left hematoma surgically removed, subarachnoid hemorrhage, ruptured Berry aneurysm	Left CVA, 9/80
General Location of Lesion	Left hemisphere	Broca's area (neurology report 9/79)	Left hemisphere	Left temporoparietal region	Left hemisphere
Time Post-Onset (at initiation of treatment	9 months	2 years	2 years	7 months	8 months

Figure 1. Visual representation of quarter-intersect results: slope, trend envelope, prediction envelope.



subject availability. Performance was measured and quantified for 28 consecutive treatment sessions. In order to view and evaluate progress, these percentage scores were plotted in real time using a separate graphic display for each task. An example display is shown in Figure 1.

Other Measurements

Trend envelopes and prediction envelopes were drawn for each data set, as shown for hypothetical data in Figure 1. The *trend envelope* measure was derived from the first half of each data set (first 14 measurements) and is described as two accelerating parallel lines that define the performance range and include all data in the initial half of the data set. A *prediction envelope* is an extension of the trend envelope (Figure 1), and it predicts a probable range of performance if progress were to continue with the same rate and variability as in the trend envelope. Specific procedures for determining trend and prediction envelopes have been described in detail previously by Matthews and LaPointe (1981).

Data Analysis

Slope Measurements

The quarter-intersect procedure has been described by Matthews and LaPointe (1981) as a potentially useful tool for rapidly evaluating the degree and rate of change that occurs during clinical intervention. The data described above (graphically displayed percentage scores) were subjected to the quarter-intersect procedure to determine the rate and direction of progress (slope) for the first 6, 10, 14, and 28 measurements for each set of performance scores. In addition, each complete set of 28 scores was subjected to the statistical procedure of least squares analysis to compare the quarter-intersect results to statistically derived slope estimations.

In the current study the consistency of each patient's progress was evaluated by two methods. First we visually compared the data displays to each corresponding prediction envelope (Figure 1). Progress was evaluated as consistent when the scores were accurately predicted by the prediction envelope, having not more than two consecutive scores outside the predicted range. The presence of a third consecutive score outside the prediction envelope indicated that the rate of progress had changed significantly and that the prediction envelope would not apply to subsequent data. According to these criteria, Figure 1 shows that (in the example) consistent progress ended with the eleventh performance score. The number of scores consistent with the predicted range (prediction limit) was determined for all sets of performance data (in this example that number is 11), and these are presented in Table 2.

A second evaluation of consistency was conducted and can be observed in Table 3. As mentioned, the slopes for the first 6, 10, 14 and 28 measurements were determined for each set of performance scores. The mathematical sign for each slope was recorded in the table. Slopes with an upward trend (positive mathematical sign) reflected a performance gain. Negative (downward) and zero (flat) slopes indicated that favorable change in behavior had not occurred. Progress was consistent if the same mathematical sign (positive, negative, or zero) was reported for the final measurement (of 28 scores) and for the early slope measurements (at 6, 10 and 14 scores).

Table 2. Number of sessions consistent with the range predicted by the prediction envelope (prediction limit).

Task	Patient	No. of Sessions Consistent with predicted Range
1	F.P.	14
2	F.P.	14
3	F.P.	14
4	F.P.	14
5	F.B.	14
6	F.B.	1*
7	F.B.	X**
8	F.B.	6*
9	R.E.	14
10	R.E.	14
11	D.M.	3*
12	D.M.	7*
13	H.E.	14
14	H.E.	14
15	H.E.	14
16	H.E.	14

* Task shift suggested
** Patient reached criterion before prediction envelope started

The mathematical sign of the final slope measurement also was used to determine whether or not measurable progress occurred during the intervention program. As mentioned in the preceding paragraph, a positive mathematical sign indicated that the patient improved his performance in the targeted skill; and negative and zero mathematical signs indicated that the expected favorable changes had not occurred. Table 3 may be useful in clarifying this point. On final slope measurement (28 treatment sessions), 14 of the 16 slopes carried positive mathematical signs.

Results and Discussion

Quarter-Intersect Compared to Least Squares

To test the accuracy of quarter-intersect trend estimations, the statistical parameters of slope and y-intercept were derived for the quarter-intersect and lease squares trend estimations. (The vertical axis [y-axis] is intercepted by the trend estimation [slope] at the y intercept. In Figure 1, y-intercept = 25%.) Correlation coefficients were determined for both types of trend estimation and for the relationship between both types of y-intercept estimations. The correlation between quarter-intersect and least-squares results was both positive and significant (slope: r^2=.96; p<0.001), (y-intercept: r^2=.708; p<0.001).

The quarter-intersect method had already proven itself to be a quick and easy method for estimating the trend of serial data (Matthews & LaPointe, 1981, 1983; White & Haring, 1976). The results of this study suggest that it also yields valid results, comparable to those derived from the complicated and time consuming least-squares statistical procedures.

Using Quarter-Intersect in Making Clinical Decisions

Early Evaluation of Tasks

Initially we asked the question: How many times does the patient need to participate in the task before we can be reasonably sure that it will help to improve his skill? To respond to this question we determined whether or not measurable progress occurred during the intervention program (as indicated by the mathematical sign) and then determined whether or not measurable progress had been occurring by the first 6, 10, and 14 sessions for the same sets of performance data. These were compared and the results can be viewed in Table 3. When the final slope measurements were taken (column 3 of Table 3), 14 of the 16 slopes carried positive mathematical signs, indicating that measurable performance gains took place. Two slopes, however, carried a neutral mathematical sign (slope equal to zero). In these cases the patient did not gain or lose skill during his treatment program. (Note that H.E. was diagnosed as having a degenerative neurological disease as a result of these findings.) In no case did a patient's measured performance deteriorate during the 28 treatment sessions, but, if it had, that slope would have carried a negative mathematical sign.

Having determined that certain tasks were ultimately beneficial to the individuals who participated and that other

Table 3. Early indications of progress compared to final measurements of whether or not progress occurred.

Task	Patient	Mathematical Sign of Final Slope Measurement	Mathematical Sign of Some Sets of Early Data: First 6 Points	First 10 Points	First 14 Points
1	F.P.	+	+*	+*	+*
2	F.P.	+	—	+*	+*
3	F.P.	+	+*	+*	+*
4	F.P.	+	0	—	+*
5	F.B.	+	+*	+*	+*
6	F.B.	+	+*	+*	+*
7	F.B.	+	+*	+*	+*
8	F.B.	+	+*	+*	+*
9	R.E.	+	0	—	—
10	R.E.	+	+*	+*	+*
11	D.M.	+	+*	+*	+*
12	D.M.	+	+*	+*	+*
13	H.E.	+	+*	+*	+*
14	H.E.	+	+*	+*	+*
15	H.E.	0**	—*	+	0*
16	H.E.	0**	+	+	0*
Ratio of Tasks with Same Mathematical Sign as Final Slope Measurement			12/16	12/16	15/16
Percent of Tasks with Same Mathematical Sign as Final Slope Measurement			75%	75%	93%

* Mathematical sign agrees with final slope measurement
** Note that H.E. was diagnosed as having a degenerative neurological disease as a result of these findings.

tasks were not, the early slope measurements (taken at 6, 10 and 14 sessions) were compared to the final measurement. We found that the direction (upward or non-upward) of the initial slope measurements (for both 6 and 10 scores) was an accurate predictor of overall outcome 75% of the time as shown in Table 3. Therefore, for these patients, six treatment sessions and six corresponding scores provided the clinician with sufficient data to make a data-based decision on task termination or continuation, if the standard of 75% accuracy is acceptable. When we included performance data for 14 sessions, accuracy of prediction was increased to 93%. These data indicate that evaluating the productivity of a treatment program prior to collecting a minimum of six consecutive performance scores may give results that are accurate less than 75% of the time. Unfortunately, many speech-language pathologists are required to make clinical decisions regarding the productivity of treatment prior to six treatment sessions; and we are rarely allowed up to 14 sessions to determine whether or not a patient is showing progress in his treatment program. These data also indicate that continuing with a task beyond 14 sessions when measurable gains have not yet occurred may not be an appropriate clinical activity.

Tasks whose slopes do not carry a positive mathematical sign for 6 to 14 treatment sessions should be altered, supplemented, or discarded and replaced by tasks that produce a positive treatment effect. (Alteration of task difficulty can be done by selecting less complex stimuli, adding prompts or cues, or allowing for greater response latency.) We assume that discarding unproductive tasks and replacing them with positively accelerating tasks will lead to a reduction in patient frustration.

Treatment Gains May Slow Down

For each task, the number of scores consistent with the predicted range are presented for each set of scores in Table 2. Subjects maintained their initial rates of progress on 11 of the 16 tasks. Conversely, their rate of progress decreased on four of the tasks. Perhaps a rate of progress that violates the predicted range indicates that a more difficult task is needed to allow the patient to progress at earlier steeper rates of change. When a subject's scores dropped below the predicted range he may have been ready to attempt more difficult tasks in addition to or in lieu of the ongoing treatment tasks.

Summary and Conclusions

The results of this study suggest that the quarter-intersect method (Matthews & LaPointe, 1981, 1983) not only offers a quick and easy procedure for estimating the trend of serial data, but also yields valid results which are comparable to those derived from the complicated and time consuming least-squares statistical procedure.

Quarter-intersect estimations may allow the clinician to ascertain early indications of progress during the initial stages of intervention. In this study, when early indications of progress were evaluated using quarter-intersect estimations, at least six consecutive treatment sessions (and corresponding performance scores) were needed to judge the measurable effect of an individual treatment task with 75% accuracy. When the procedure was applied to 14 early performance scores, predictions of whether progress would occur on that task for the patient was increased with accurate predictions occurring 93% of the time.

Quarter-intersect based slope measurements and predictions may be useful to plot rate and degree of reacquisition of communication skills. Characteristics and variability of performance rates and how these slopes interact with variation in subject and task variables remains to be determined; but the quarter-intersect procedure can be a useful tool for determining these interactions, as well as for measuring the slope and evaluating the utility of specific individualized tasks. The need for data-based clinical decisions on intervention strategies and treatment packages has been a long-standing one. Further exploration of criteria for evaluating task effectiveness, prediction of language reacquisition rates, and strategies for data-based clinical decision-making certainly will be welcome.

Acknowledgments

This project was conducted at the University of Florida and Gainesville Veteran's Administration Medical Center in conjunction with a doctoral dissertation completed under the direction of Leonard L. LaPointe, Ph.D., Thomas B. Abbott, Ph.D., G. Paul Moore, Ph.D., Warren J. Rice, Ph.D., and William D. Wolking, Ph.D. served as dissertation committee members. Randolf Carter, Ph.D. served as statistical consultant. I wish to express my sincere gratitude to all of those mentioned above.

Address all correspondence to:
Barbara Ann Johnson
Assistant Professor
National Technical Institute for the Deaf
One Lomb Memorial Drive
Rochester, NY, 14623

References

Benson, D. F. (1979). Aphasia rehabilitation. *Archives of Neurology, 36*, 187-189.

Darley, F. L. (1972). The efficacy of language rehabilitation in aphasia. *Journal of Speech and Hearing Disorders, 37*, 3.

Darley, F. L.(1982). *Aphasia.* Philadelphia: W. B. Saunders Co.

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

Matthews, B. A. Johnson, & LaPointe, L. L. (1981). Determining rate of change and predicting performance levels in aphasia therapy. *Clinical aphasiology conference proceedings.* R. Brookshire (Ed.), Minneapolis: BRK Publishers.

Matthews, B. A. Johnson, & LaPointe, L. L. (1983). Slope and variability of performance on selected aphasia treatment tasks. In R. Brookshire (Ed.), *Clinical aphasiology conference proceedings.* Minneapolis: BRK Publishers.

Porch, B. E. (1967). *Porch index of communicative abilities.* Palo Alto: Consulting Psychologists Press.

Prognosis for Aphasia. (1977). *Lancet, 2,* 24.

Sarno, M.T., Silverman, M., & Sands, E. (1970). Speech therapy and language recovery in severe aphasia. *Journal of Speech and Hearing Disorders, 13*, 607-623.

Wertz, R. T. (1986). Language disorders in adults. In J. M. Costello & A. L. Holland (Eds.), H*andbook of speech and language disorders.* San Diego: College Hill Press.

White, O. R., & Haring, N.G. (1976). *Exceptional teaching: A multi-media training package.* Columbus, OH: Charles E. Merrill.