CHAPTER

5

Analytical and Technical Directions in Applied Aphasia Analysis: The Midas Touch

Cynthia K. Thompson
Kevin P. Kearns

Over a decade ago, clinical aphasiologists began discussing the application of applied behavior analysis techniques to the study of treatment for aphasia (Davis, 1978a, 1978b; LaPointe, 1978). Recognizing the need for "our data not our word" (Davis, 1978b) in documenting the effects of treatment, single-subject experimental designs began to appear in the literature. The purpose of this paper is to present findings from a review of treatment studies published in *Clinical Aphasiology* during the past 10 years, from 1978 to 1987, in which single-subject designs were utilized. The review was undertaken not only to delineate trends pertaining to treatment efficacy research from an historical perspective, but also to address future directions of applied aphasia research.

Treatment studies reviewed were identified by two independent judges rating all articles published in *Clinical Aphasiology* during the years of interest. As discussed by Kearns and Thompson in Chapter 4, a relatively consistent number of articles concerned with treatment for aphasia were found each year; however, an increase in the use of single-subject experimental designs was evident along with a concomitant decrease in case studies.

Of particular interest in our review was to address aspects of applied experimental analysis discussed by Baer, Wolf, and Risley (1968) in a seminal article entitled, "Some Current Dimensions of Applied Behavior Analysis." Dimensions selected for consideration in the present review were (1) conceptual, (2) analytical, and (3) technical. Analytical and technical aspects are discussed in the present paper, whereas conceptual aspects are addressed in Chapter 4.

METHOD

ANALYTICAL DIMENSIONS

Articles utilizing single-subject experimental designs were rated on three dimensions of analysis in treatment research: *effectiveness*, *efficiency*, and *effects* (Kendall and Norton-Ford, 1982; Olswang, in press; Rosen & Proctor, 1981; Thompson, in press). In evaluating these dimensions, the type of single-subject design employed was coded and experimental questions asked were noted.

Treatment Effectiveness

Effectiveness studies, concerned with demonstration of experimental manipulations, that is, cause-effect relationships between independent and dependent variables, were coded as such when experimental ques-

tions such as the following were posed: *"Is treatment effective in changing aphasic behavior?"* (Hays, Rincover, and Solnick, 1980). Studies addressing such questions utilized the familiar withdrawal and reversal designs [ABA(B) and its variations] or the various multiple-baseline designs (Barlow and Hersen, 1984; Kearns, 1986; McReynolds and Kearns, 1983).

Treatment Efficiency

Experiments in the literature coded as being efficiency evaluations were concerned with (1) examining the relative effectiveness of treatment variables by comparing two or more treatment approaches, (2) determining what components of treatment packages are most important for changing aspects of aphasia, and (3) evaluating the intensity or frequency of a particular independent variable that is most efficacious. Studies examining the relative effectiveness of treatments required an alternating treatment design (ATD), an ABACA with counterbalancing, or a replicated crossover design (Barlow and Hersen, 1984; Hegde, 1987; Kearns, 1986; McReynolds and Kearns, 1983). To evaluate the relative contribution of treatment components, an interaction design [interaction additive or interaction reduction (Kearns, 1986)] was required in which treatment variables were systematically added or deleted throughout the course of treatment. Finally, examination of variables such as the intensity, frequency, or other parameters of treatment required a changing-criterion design.

Treatment Effects

Studies were coded as being concerned with treatment effects when generalization was examined or when social validation data were obtained. Evaluation of these outcomes could be accomplished using several of the single-subject designs; therefore, the focus of analysis in this area was concerned with whether or not questions addressing these issues were asked and whether or not measures were taken to answer them. Researchers addressing generalization posed questions to determine when, if ever, in the treatment process a patient began to progress or change in areas not directly treated (i.e., across behaviors or settings). Studies concerned with social validity addressed functional outcomes of treatment through the gathering of social comparison or subjective evaluation data. Questions of this nature included the following: *"Will treatment result in a patient communicating more successfully or 'normally' in functional situations?"* or *"Will the spouse of the aphasic patient perceive improvement in communication following treatment?"*

Five types of generalization were individually scored—generalization across behaviors, time (maintenance), stimulus conditions (when untrained task stimuli were used to elicit trained responses), settings, and persons. Each study was scored as to whether or not each type of generalization was experimentally assessed, independent of whether it was achieved. Generalization across time was coded for both within-session testing (as in the multiple-baseline design when early trained behaviors continue to be probed while new behaviors are trained) and follow-up maintenance testing, occurring at least 1 week removed from the termination of treatment (Thompson, 1988). Studies were coded with regard to social validation by evaluating each paper for the inclusion of social comparison or subjective evaluation data (Goldstein, in press; Kazdin, 1977; Thompson, 1988).

TECHNICAL DIMENSIONS

Six technical dimensions were evaluated. These included intrasubject and intersubject replication, counterbalancing to control for order effects, reliability on independent and dependent variables, and internal validity. Of concern in this analysis was determining to what extent these technical components of single-subject research are included in applied studies of aphasia.

INTEROBSERVER RELIABILITY

Each article was read and independently scored on 16 separate dimensions (3 analytical dimensions, 7 in the effects category, and 6 technical dimensions) by both authors. Reliabilities ranged from 66 to 100 percent across the 16 categories, with overall reliability at 91 percent.

RESULTS

ANALYTICAL DIMENSIONS

Effectiveness

Effectiveness studies—in which the form of the experimental question was simply, *"Was there an effect of B?"*—comprised the majority of the

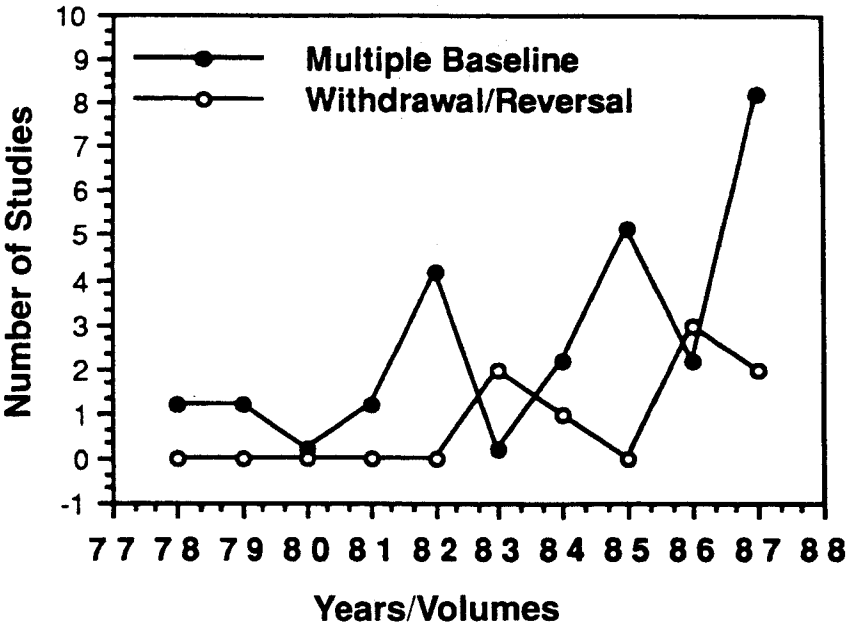

Fig. 5-1. The number of multiple-baseline and withdrawal or reversal (ABA(B)) designs used to investigate experimental questions concerned with the *effectiveness of treatment* in *Clinical Aphasiology* from 1978 to 1987.

applied experimental work in *Clinical Aphasiology*. Of a total of 33 applied studies identified, 28, or 85 percent, addressed questions of this nature using either a multiple-baseline or an ABA(B) design. Figure 5-1 indicates that increases in both multiple-baseline designs and withdrawal or reversals were found over the past 10 years, with the most frequently used method of analysis being the multiple-baseline design. These trends were not surprising, since it is expected that in early stages of validating treatment efficacy most studies will be concerned with establishing the effectiveness of treatment variables rather than with comparing relative efficiency or evaluating components of treatment (Baer, Wolf, and Risley, 1968). As pointed out by Hayes, Rincover, and Solnick (1980), "any type of experimental analysis must first be based on establishing overall effectiveness. One can hardly ask about the effects of components, for example, until the overall effectiveness is clear" (p. 277).

Efficiency

As suggested by the number of effectiveness studies, few efficiency studies were identified. Studies in which such questions were posed as, *"Is B more (or less) effective than C?" "What components of B make it an effective treatment?"* or *"What are the effects of different amounts (levels) of B?"*

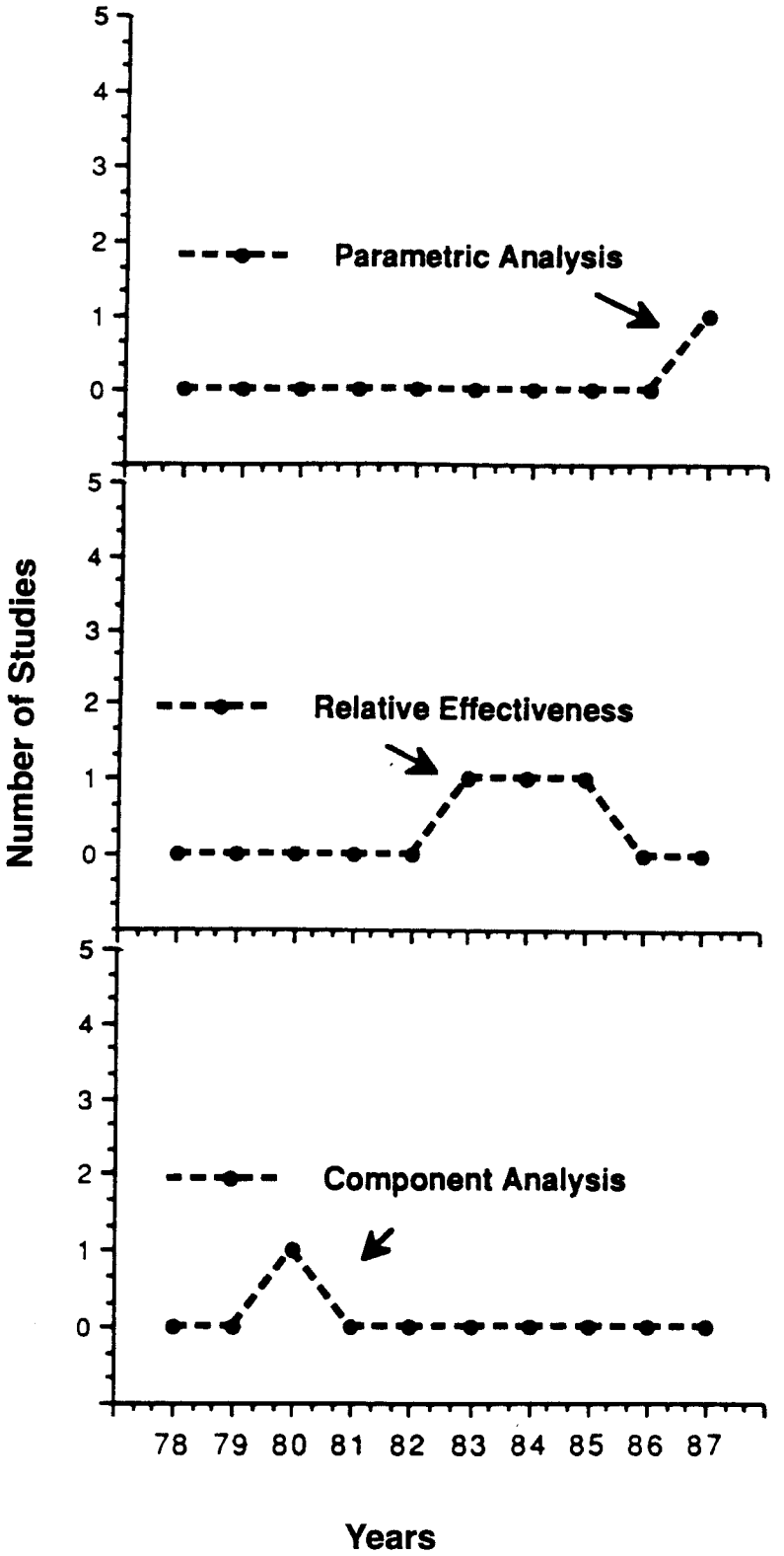

Fig. 5-2. The number of experimental studies evaluating the *efficiency of treatment* through component analysis, analysis of relative effectiveness of treatments, and parametric analysis in *Clinical Aphasiology* from 1978 to 1987.

were not common in *Clinical Aphasiology*. A total of only five studies in 10 years were found (see Fig. 5-2), comprising 15 percent of applied aphasia research. Specifically, one component analysis (Simmons, 1980), three relative effectiveness studies (Burger and Wertz, 1984; Golper and Rau, 1983; Loverso et al., 1985), and one parametric analysis (Southwood, 1987) were found. These findings suggested that increased studies specifying types and aspects of treatment for aphasia that are most efficient are needed. While efficiency studies are not superior to effectiveness studies, they are a natural outgrowth of or follow-up to studies of effectiveness. For example, once a particular treatment (B) is found to be effective, a next step in programmatic research might be to determine its efficiency, for example, by comparing B to other treatments, or by dissecting components of B and evaluating their effectiveness, or by introducing B in increasingly larger doses. The lack of such research supports the impressions of Kearns and Thompson (Chap. 4) that increased programmatic research is needed in *Clinical Aphasiology*.

Effects

Review of the final analytical dimension, the effects of treatment, indicated trends toward increased attention to generalization but not to social validation of treatment effects. With regard to generalization, Figure 5-3 indicates that increases in the number of studies measuring generalization across behaviors, time (both within-session and follow-up maintenance), and stimulus conditions were noted, but that few studies addressed generalization across settings or persons. These findings indicated a need for inclusion of generalization testing across environments and people in future studies. Caution is warranted in interpreting the increased trends found in measurement of generalization across behaviors and time. That is, these findings may be an artifact of the use of multiple-baseline designs across behaviors in that use of the design itself requires measurement of behavior prior to training and testing of maintenance following training. If a more strict definition of generalization had been used, as per Hayes, Rincover, and Solnick (1980), who only consider studies to be generalization experiments when measures are taken with no planned intervention, few studies measuring generalization across behaviors and/or within-session maintenance in the present review would have been identified.

In addition, few studies reported social validation data. That is, data demonstrating that aphasic patients communicate better following treatment and that people who communicate with them think that they do are lacking in *Clinical Aphasiology*. As seen in Table 5-1, only four studies in 10 years have reported data of this nature that were collected systematically and reliably.

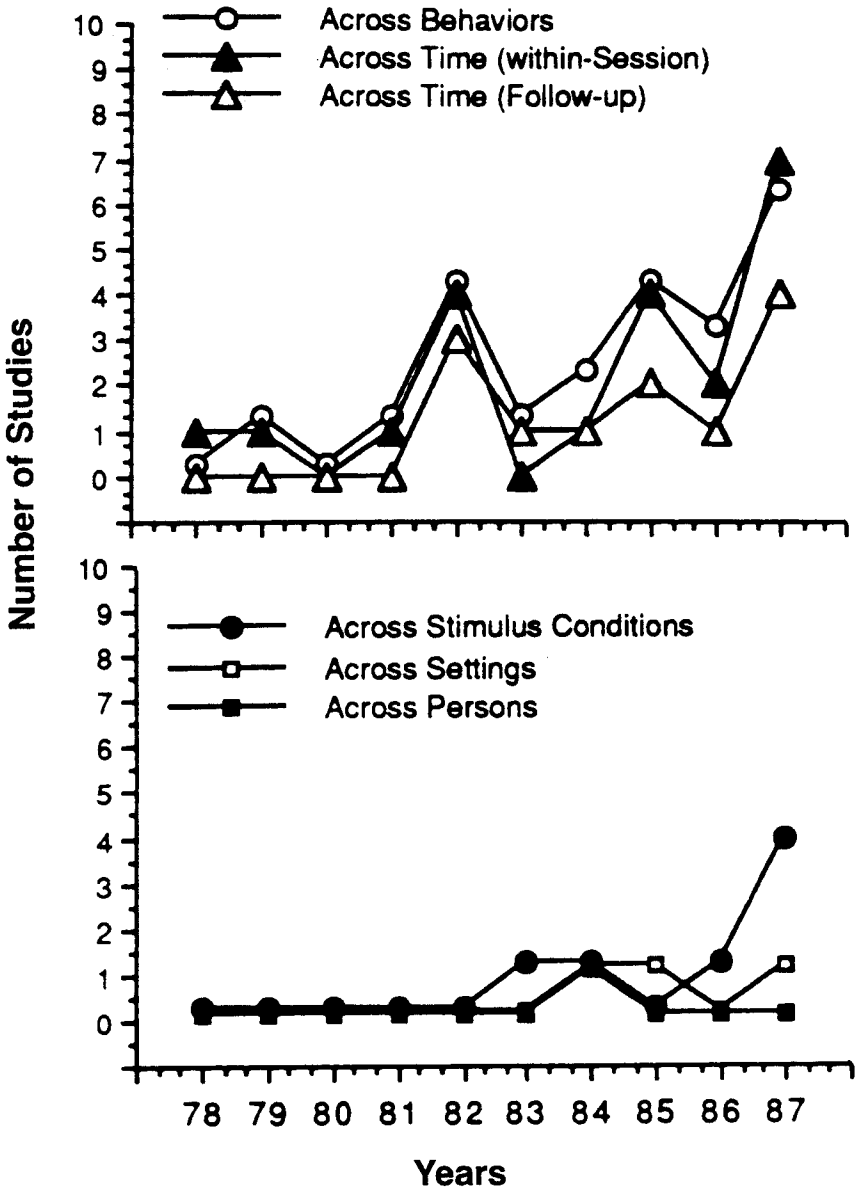

Fig. 5-3. The number of experimental studies addressing the *effects of treatment* through assessment of generalization across behaviors, time, stimulus conditions, settings, and persons in *Clinical Aphasiology* from 1978 to 1987.

TECHNICAL DIMENSIONS

Results of the literature review with regard to technical trends are depicted in Table 5-2. The total number of applied studies reported each year are shown as well as the number of studies accomplishing replication both within (intrasubject) and across subjects (intersubject), counterbalancing of variables, reliability on the dependent and independent variables, and internal validity.

TABLE 5-1. NUMBER OF EXPERIMENTAL STUDIES REPORTING SOCIAL VALIDATION DATA FROM 1978 TO 1987

Year	*Total number of experimental studies*	*Number reporting social validation data*
1978	1	0
1979	1	0
1980	1	0
1981	1	0
1982	4	0
1983	3	0
1984	3	1
1985	5	1
1986	4	0
1987	10	2
TOTAL	33	4

Replication and Counterbalancing

Intrasubject replication, or studies in which a treatment effect was replicated within subjects, was demonstrated in 0 percent of studies reported in 1978 (0 of 1) and in 80 percent of studies reported in 1987 (8 of 10), with an overall mean of 66 percent. A similar increase over the years, however, was not noted with regard to intersubject replication. Of a total of 33 studies, only 33 percent (11 of 33) demonstrated treatment effects on more than one subject.

It is well known that single-subject research is not synonymous with studying only one subject (McReynolds and Thompson, 1986), and the replication requirement in single-subject research for addressing the lawfulness of findings and external validity is well documented (Barlow and Hersen, 1984; Sidman, 1960). The common reporting of data from only one subject in *Clinical Aphasiology* is perhaps an oversight among clinical aphasiologists.

Analysis of the data indicated that counterbalancing to control for order effects also seldom is accomplished, with only 18 percent (6 of 33) of studies doing so over the years. Counterbalancing is particularly important in multiple-baseline studies across behaviors because the design requires sequential applications of treatment across behaviors and training of early behaviors may influence the training of later ones. Because multiple subjects are required to counterbalance variables, the noted

TABLE 5-2. TECHNICAL TRENDS IN EXPERIMENTAL STUDIES REPORTED BETWEEN 1978 AND 1987

		Subject replication			*Reliability**		
Year	*No. of studies*	*Intra*	*Inter*	*Counterbalancing*	*Dep. v.*	*Ind. v.*	*Internal validity*
1978	1	0	0	0	0	0	0
1979	1	0	1	0	0	0	0
1980	1	1	0	0	0	0	1
1981	1	1	0	0	0	0	1
1982	4	3	3	2	4	0	3
1983	3	1	0	0	0	0	0
1984	3	1	1	0	2	0	2
1985	5	4	3	3	5	0	5
1986	4	3	1	0	3	0	4
1987	10	8	2	1	7	0	6
Total	33	22	11	6	21	0	22
Percent		66	33	18	63	0	66

*Dep. v. = dependent variable; Ind. v. = independent variable.

lack of counterbalancing perhaps is related to the observation that more than one subject is seldom studied.

Reliability

Interobserver agreement or reliability was reported on dependent variables in 0 percent of studies in 1978 and in 70 percent of studies in 1987, with an overall mean across years of 63 percent. No studies in the 10-year period, however, reported interobserver reliability on the independent variable(s). As reasoned by Yeaton and Sechrest (1981), McReynolds and Thompson (1986), and others, if changes in behavior are to be explained by treatment variables, it is important that experimenters provide evidence that treatment was administered as described.

Internal Validity

Finally, in terms of internal validity, a low of 0 percent of studies in 1978 demonstrated experimental control, while a high of 100 percent demonstrated experimental control in 1986. The overall mean across years was 66 percent, leaving a remaining 44 percent of single-subject studies in the aphasia literature uncontrolled. This finding indicates that the mere use of single-subject designs does not guarantee that experimental control will be forthcoming. For example, an ABA(B) design without a demonstrated reversal in the second A phase (e.g., Berstein-Ellis, Wertz, and Shubityowski, 1987) is in reality an AB design (case study) with maintenance testing in the final A phase.

SUMMARY AND DISCUSSION

In summary, findings from this review indicate an increase in the use of single-subject designs to validate treatment effects. However, most applied aphasia studies focused on evaluating effectiveness, with few investigating treatment efficiency. Increased concern with generalization across behaviors and time but not across settings and persons was noted, and social validation of outcome data was not common. Technically, increased intrasubject but not intersubject replication was demonstrated over the years, increased concern with observer reliability on dependent variables but not on independent variables was noted, and finally, increased demonstrations of internal validity were prevalent in the literature.

Overall, these findings indicate that efforts in the area of applied aphasia analysis have accelerated over the past decade. A change in focus from case studies to controlled experimental studies is clear, and greater adherence to technological aspects is evident. However, these trends may be, in part, only illusive of improved science. This review points out needs for increased technological sophistication and programmatic efforts addressing efficiency and the overall effects of treatment. Efforts need to be made to examine which treatment variables are most effective and for whom, and at the same time, further efforts need to be made to validate laboratory procedures in an ecologic sense. With applied experimental analysis techniques, the effectiveness and efficiency of treatment can be evaluated, and by incorporation of generalization measures, social comparison data, and subjective evaluations of outcome, the broader effects of treatment may be determined.

One obstacle to examining these larger effects of treatment is the phenomenon known as *reductionism* (Thompson, in press; Warren, 1979). This phenomenon is central to natural science and determinism or the search for cause-effect relationships. That is, in order to study and ultimately understand a phenomenon, it is necessary to reduce it to measurable and observable levels. Because communication encompasses numbers of variables, studying it experimentally is necessarily reductionistic. However, the limitations of reductionist data need to be recognized. As experimentalists who enunciate lawful principles in our research designs, we are obliged to state the boundaries that limit the generalization of findings and to systematically and programmatically build on reductionistic data until aspects of treatment are fully understood.

The technology to advance our science is presently available, and while we are still learning how to best apply the technology to the study of aphasia, this may be considered a small problem related to the larger problem of determining how much and what kind of treatment is best and what changes constitute important treatment outcomes. Perhaps at this point treatment research in aphasia has become only gilded with gold through the use of single-subject designs. What is needed now is not different than what was needed prior to this trend, when case studies outnumbered experimental studies. That is, a closer look at substances underneath, such as the influence of organism variables (e.g., patterns of language deficit) on treatment outcomes, along with a building of increased layers of information regarding the type and amount of treatment needed to change and improve observed deficit patterns, is needed. These activities, along with increased design sophistication, will constitute the midas touch.

REFERENCES

Barlow, D. H., and Hersen, M. (1984). *Single case experimental designs: Strategies for studying behavioral change,* 2d Ed. New York: Pergamon Press.

Baer, D. M., Wolf, M. M., and Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1,* 91–97.

Bernstein-Ellis, E., Wertz, R. T., and Shubitowski, Y. (1987). More pace, less fillers: A verbal strategy for a high-level aphasic patient. In R. H. Brookshire (Ed.), *Clinical aphasiology,* Vol. 17 (pp. 12–22). Minneapolis, MN: BRK Publishers.

Burger, L. H., and Wertz, R. T. (1984). The effect of a token training program on auditory comprehension in a case of Wernicke's aphasia. In R. H. Brookshire (Ed.), *Clinical aphasiology,* Vol. 14 (pp. 173–180). Minneapolis, MN: BRK Publishers.

Davis, G. A. (1978a). The clinical application of withdrawal, single-case research designs. In R. H. Brookshire (Ed.), *Clinical aphasiology,* Vol. 14 (pp. 11–19). Minneapolis, MN: BRK Publishers.

Davis, G. A. (1978b). Discussion: "Our data not our word." In R. H. Brookshire (Ed.), *Clinical aphasiology,* Vol. 8 (pp. 37–39). Minneapolis, MN: BRK Publishers.

Goldstein, H. (in press). Clinical significance. In L. B. Olswang, C. K. Thompson, S. Warren, and N. Minghetti (Eds.), *Treatment efficacy research in communication disorders.* Rockville, MD: American Speech-Language-Hearing Foundation.

Golper, L., and Rau, M. (1983). Systematic analysis of cuing strategies in aphasia: Taking your "cue" from the patient. In R. H. Brookshire (Ed.), *Clinical aphasiology,* Vol. 13 (pp. 52–61). Minneapolis, MN: BRK Publishers.

Hayes, S. C., Rincover, A., and Solnick, J. V. (1980). The technical drift of applied behavior analysis. *Journal of Applied Behavior Analysis, 13,* 275–285.

Hegde, M. N. (1987). *Clinical research in communicative disorders: Principles and strategies.* Austin, TX: PRO-ED.

Kazdin, A. E. (1977). Assessing the clinical or applied significance of behavior change through social validation. *Behavior Modification, 1,* 427–452.

Kearns, K. P. (1986). Flexibility of single-subject experimental designs: II. Design selection and arrangement of experimental phases. *Journal of Speech and Hearing Disorders, 51,* 204–214.

Kearns, K. P., and Thompson, C. K. (in press). Technical drift and conceptual myopia: The Merlin effect. In T. E. Prescott (Ed.), *Clinical aphasiology,* Vol. 19 (pp. 31–40). Austin, TX: PRO-ED.

Kendall, P., and Norton-Ford, J. (1982). Therapy outcome research methods. In P. Kendall and J. Butcher (Eds.), *Handbook of research methods in clinical psychology* (pp. 429–460). New York: Wiley.

La Pointe, L. L. (1978). Multiple baseline designs. In R. H. Brookshire (Ed.), *Clinical aphasiology,* Vol. 8 (pp. 20–29). Minneapolis, MN: BRK Publishers.

Loverso, F. L., Prescott, T. E., Selinger, M., Wheeler, K. M., and Smith, R. D. (1985). The application of microcomputers for the treatment of aphasic adults. In R. H. Brookshire (Ed.), *Clinical aphasiology,* Vol. 15 (pp. 189–195). Minneapolis, MN: BRK Publishers.

McReynolds, L. V., and Kearns, K. P. (1983). *Single-subject experimental designs in communicative disorders.* Austin, Texas: PRO-ED.

McReynolds, L. V., and Thompson, C. K. (1986). Flexibility of single-subject experimental designs: I. Review of the basics of single-subject designs. *Journal of Speech and Hearing Disorders, 51,* 194–203.

Olswang, L. B. (in press). Treatment efficacy: The breadth of research. In L. B. Olswang, C. K. Thompson, S. Warren, and N. Minghetti (Eds.), *Treatment efficacy research in communication disorders.* Rockville, MD: American-Speech-Language-Hearing Foundation.

Rosen, A., and Proctor, E. (1981). Distinctions between treatment outcomes and their implications for treatment evaluation. *Journal of Consulting and Clinical Psychology, 49,* 418–425.

Sidman, M. (1960). *Tactics of scientific research.* New York: Basic Books.

Simmons, N. N. (1980). Choice of stimulus modes in treating apraxia of speech: A case study. In R. H. Brookshire (Ed.), *Clinical aphasiology,* Vol. 10 (pp. 302–307). Minneapolis, MN: BRK Publishers.

Southwood, H. (1987). The use of prolonged speech in the treatment of apraxia of speech. In R. H. Brookshire (Ed.), *Clinical aphasiology,* Vol. 17 (pp. 277–287). Minneapolis, MN: BRK Publishers.

Thompson, C. K. (1988). Generalization in the treatment of aphasia. In L. V. McReynolds and J. Spradlin (Eds.), *Generalization strategies in the treatment of communication disorders* (pp. 82–115). Philadelphia: B. C. Decker.

Thompson, C. K. (in press). Issues in conducting treatment research: Comments on Doyle et al., Bourgeois, and Lansing and Davis papers. In L. B. Olswang, C. K. Thompson, S. Warren, and N. Minghetti (Eds.), *Treatment efficacy research in communication disorders.* Rockville, MD: American Speech-Language-Hearing Foundation.

Warren, R. L. (1979). Treatment: Introductory remarks. In R. H. Brookshire (Ed.), *Clinical aphasiology,* Vol. 9 (pp. 151–153). Minneapolis, MN: BRK Publishers.

Yeaton, W. H., and Sechrest, L. (1981). Critical dimensions in the choice and maintenance of successful treatments: Strength, integrity, and effectiveness. *Journal of Consulting and Clinical Psychology, 49,* 156–167.