Once Is Not Enough: Some Thoughts on Replication in Functional Analysis Research in Aphasia

John D. Tonkovich
Veterans Administration Medical Center, Decatur, Georgia

Richard K. Peach
West Paces Ferry Hospital, Atlanta, Georgia

In recent years, there has been growing interest in functional analysis or within-subject research designs among clinical aphasiologists. These designs developed largely as an outgrowth of the behavioral movement of the 1960s. Clinicians in applied behavioral fields such as psychology and special education desired research methodologies which could aid in the evaluation and identification of factors which could be used to alter a selected target behavior. The findings of such investigations would be utilized in the development of more efficient, effective training programs.

The merits of functional analysis research have been discussed extensively in the literature (Hersen and Barlow, 1976; Davis, 1978; Kazdin, 1982; McReynolds and Kearns, 1983). Among the advantages cited have been the following:

1. These designs permit clinician-investigators to explore relationships between treatment variables and behavior change.
2. These designs allow investigation of the effects of certain treatment variables on populations with specific disorders.
3. Small numbers of individuals can be subjected to research methodologies within a clinical context.
4. With respect to behavioral change subsequent to treatment, individual differences among subjects are highlighted by functional analysis research.
5. Functional analysis designs focus on the observation of behavioral changes over time, providing information about rate of behavioral change.

Because of these advantages, within-subjects designs provide clinician-investigators with information about the effectiveness and efficiency of treatment procedures, and may form the basis for refinements in and development of future treatment protocols. It is not surprising, given the relative newness of this type of investigation, that there exist only a few aphasia treatment studies in the past several years which employ within-subject experimental designs.

While functional analysis research designs do not typically employ statistical procedures, they are not without sound theoretical and scientific bases. Because a small sample size is frequently utilized in functional analysis designs, one criticism often raised is that the generalization of such experimental findings is limited. This is a valid criticism, and in order for findings of functional analysis research to be generalized to larger populations, the issue of replication becomes a consideration for the aphasia clinician-investigator.

All within-subject research designs have, by definition, some form of intrinsic replication. Replication, in a general way, refers to repetition or reproduction of the independent variable conditions of an experiment under different circumstances. Replication does not refer to a reproduction of the

original findings of an experiment, but rather to the extent to which similar stimulus conditions associated with earlier observations are duplicated. In fact, as Johnston and Pennypacker (1980) noted, even replications which fail to generate the same effects on behavior may be considered successful, particularly when an alternative condition is sufficient to produce the effect. The utility of replication procedures lies in their providing clinician-investigators with information about 1) the reliability or stability of a behavioral effect, and 2) the generality of a functional relation over a wide range of conditions. Both types of information are crucial when treatment effectiveness is being evaluated.

Sidman (1960) proposed two broad classes of replication techniques: direct replication, which is related to the reliability of treatment procedures; and systematic replication, which yields information relative to the generality of experimental procedures. While these two classes of replication are conceptually different, it is often difficult to classify a replication procedure into one or the other category. As McReynolds and Kearns (1983) pointed out, clinician-investigators may primarily be interested in direct replication, when circumstances force systematic replication to occur.

Direct replication, in its simplest form, is the repetition of a given experiment by the same investigator. This is accomplished by repeating the experiment with new subjects (intersubject replication) or by repeating the experiment on the same subjects (intrasubject replication). According to Hersen and Barlow (1976), implicit in the notion of intersubject replication are the assumptions that subjects are homogeneous, the experimenter remains constant, and the experimental setting remains the same.

Intersubject replication in the aphasia literature is exemplified in a study by Kearns and Salmon (1983, in press). In this investigation, two individuals with longstanding Broca's aphasia subsequent to trauma were subjected to the same training by the same experimenter. The study was undertaken to determine whether training auxiliary _is_ forms would generalize to untrained auxiliary and copular _is_ forms in other contexts. The two aphasic subjects performed similarly. It should be noted that Hersen and Barlow (1976) advocated three replications (a total of four subjects) as being sufficient for establishing reliability of treatment procedures, and the Kearns and Salmon study highlights one of the problems clinical aphasiologists face in performing intersubject replications. It is not often easy to meet the homogeneity of subjects criterion for such replications. Clinicians might have to wait a number of years before having access to four individuals with "homogeneous" aphasias. Addressing this issue, McReynolds and Kearns (1983) offered some compensatory alternatives, such as obtaining exhaustive baseline measures, conducting the study anyway, noting differences among subjects, and searching for the source or sources of variation, when subject-matching becomes difficult or impossible.

While intersubject replications provide clinician-investigators with some hypothetical notion of the range of subject variation allowable for predicted treatment success, intrasubject replications provide preliminary evidence that the obtained functional relationship did not occur randomly. Intrasubject replication is typified as well in the Kearns and Salmon (1983) study. An A-B-A-B reversal design was employed. In a general way the second A-B pair (in this case, reversal to baseline levels followed by auxiliary _is_ retrainings) provided replication of the first A-B pair (baseline, then auxiliary _is_ training). For each of the two aphasic individuals in this study, the data suggested greater accuracy levels during the training phases, although the rates of change were somewhat disparate within subjects.

Once direct replications have been performed, the clinician-investigator can generate a number of questions relative to subsequent replication attempts. In order to provide answers to these questions and ultimately to permit generalization of findings to other individuals, behaviors, clinicians, and/or settings, systematic replication experiments need to be performed. According to Sidman (1960):

> "If systematic replication fails, the original experiment will still have to be redone...[because] there is no way of determining whether the failure to replicate stemmed from the introduction of new variables in the... [systematic replication] experiment, or whether the control of relevant factors was inadequate...if systematic replication succeeds, the pay-off is handsome. Not only is the reliability of the original finding increased, but also its generality with respect to other organisms and to other experimental procedures is greatly enhanced." (pp. 111-112)

Several systematic replications have occurred in the recent aphasia treatment literature (Sanders, Davis, and Hubler, 1979; Thompson, McReynolds, and Vance, 1982). Two studies will be discussed relative to their potency as systematic replications.

I would like to comment first of all on a paper I presented at the Clinical Aphasiology Conference last year (Tonkovich and Loverso, 1982). In that experiment, we demonstrated that four nonfluent aphasic patients were able to generalize trained verb + object gesture combinations to untrained verb + object pairs using matrix training procedures. These findings replicated earlier findings from the child language literature. It should be noted that the subjects in this study were highly disparate in terms of severity and length post onset of aphasia, and that two different clinicians in different clinical settings administered the treatment. On the surface, these data lend strong support to the notion of generalizing the methodology to other nonfluent aphasic patients in other contexts. This discovery, in fact, was more serendipitous than planned. Had the patients' performances on the experimental task differed widely, these data would have been collected in vain, as no previous direct replication experiments had been performed to establish the reliability of the procedures.

The notion of entering into a within-subject experiment blindly is not unique to our study, and not uncommon in the aphasia treatment literature using within-subject designs. Several factors may account for this observation. McCarthy (1983) speculated that replicative works are not as highly regarded in the scientific community as those demonstrating creative thinking. Therefore, aphasia treatment investigators may generate original ideas and consequently undertake functional analysis experiments, overlooking the theoretical underpinnings of these research designs--a "treat and hope" philosophy. A second factor may be that aphasia clinician-investigators, eager to perform experiments with the relatively new within-subject designs, engage in them, although they are not fully cognizant of critical design considerations. An experiment is typically completed and viewed as an endpoint, rather than one step in a longitudinal series of experiments. If clinical aphasiologists are going to be able to use within-subject methodologies to demonstrate treatment effectiveness, they must develop a longitudinal scope of research goals, and use the data from each within-subject experiment as a starting point for generating subsequent research questions. Systematic replication should be planned and controlled, rather than accidental.

A study in the aphasia literature, which perhaps best exemplifies an overt attempt at systematic replication is one by Stanton, Yorkston, Aune,

and Hedges (1982). These investigators attempted to replicate systematically a study by Tonkovich and Berman (1981) to demonstrate that the two-alternative forced-choice paradigm could be used to train patient error recognition for written language deficits. In the original study, training was provided to an individual with fluent aphasia secondary to vascular causes. The Stanton et al. (1982) investigation replicated the methodology with a head injured patient--a systematic replication across populations, clinicians, and settings. The Stanton et al. (1982) data supported the findings of the earlier study and can support the utility of error recognition training for written language disturbances, and extend the application of the procedure to a different clinical population.

A third category of replication procedures (proposed by Hersen and Barlow, 1976) is clinical replication. Clinical replication is "an advanced replication procedure in which a treatment 'package' containing two or more distinctive procedures is provided to a succession of clients with multiple behaviors [which] cluster together." Hersen and Barlow, 1976, p. 335) A clinical replication can occur only after extensive, controlled experimental studies yield data which isolate effective, potent treatment variables. These treatment variables are combined into a treatment package of two or more procedures, and the clinical replication serves to establish the potency of the package to effect behavioral change. Clinical replication is rare in the behavioral analysis literature, and has not been attempted in the aphasia treatment literature. The major reason for this with respect to the aphasia literature is that no treatment programs have been developed as a result of controlled, systematic investigation.

While we do have some protocols for treating aphasia which are quite explicit and fairly rigorously defined (such as Melodic Intonation Therapy and Visual Action Therapy), the elements of these methodologies have never been subjected to rigorous experimental investigation. Consider, for instance, that several steps in Melodic Intonation Therapy might be eliminated without radically altering the ultimate outcome of the treatment package. We will never know whether MIT can be provided more efficiently until the individual treatment variables of this methodology are subjected to experimental analysis. While programs such as MIT and VAT have proven to be useful agents for effecting behavioral changes in our clinical environments, they may not be the most efficient procedures. As McReynolds and Kearns (1983) pointed out, clinicians can place more confidence in programs derived from scientifically obtained data than from programs developed solely on hunches and clinical intuition.

Clinical replication, then, provides strong support for the notion of clinical relevance and acceptability of our treatment procedures. Closely tied to clinical relevance is the concept of social validation. In a general way, social validity refers to subjective judgments by society about the acceptability of objective behavioral research. According to Wolf (1978), there is a need for social validation of behavioral research on at least three levels:

1. The social importance of the goals. Are the specific behavioral goals what society really wants?
2. The social appropriateness of the procedures. Do participants and other consumers consider the treatment procedures acceptable?
3. The social importance of the effects. Are consumers satisfied with the results (including any unpredicted side-effects) of treatment?

There is an increasing trend in the recent behavioral analysis literature to validate experimental findings from the perspective of the impact of the

findings on society. This is a notion that is rarely considered in the aphasia literature.

Recently Post (1983), in a personal account of her recovery from aphasia, confessed that when the speech and language pathologist came to provide treatment, she pretended to be asleep rather than face the treatment she hated. We probably rarely, if ever, ask our patients whether they like pointing to pictures of objects when we provide them with the function, or whether it makes any difference in their daily lives. Coupled with the recent surge of interest about the pragmatic aspects of language, the concept of social validation may prove to be an important means of establishing that our treatment effects do generalize to daily living situations, and that these effects can be recognizable to others. The aphasic patient, in a sense, is at the mercy of the speech and language pathologist when it comes to determining the course of speech and language intervention procedures and goals--aphasic patients, for the most part, tend to do as we tell them. Perhaps our awareness of the need for social validation with regard to behavioral aphasia treatment can lead to more effective treatment on a social plane as well.

In summary, the merits of functional analysis or within-subject experimental designs for evaluating the efficacy of aphasia treatment are obvious. Replication procedures are an integral component of demonstrating treatment effectiveness, and should not be regarded lightly by the clinical aphasiologist. In aphasia intervention experiments, there is need for direct replication--data from both intersubject and intrasubject replications--to aid in establishing the reliability of treatment procedures. Systematic replications across subjects, clinicians, and settings is needed to provide generality evidence about the treatment procedures. There is need for development of treatment packages which arise from solid experimental data and a need to subject these treatment packages to rigorous scrutiny via clinical replications. Finally, if our treatment procedures are effective in the clinical setting, we need to investigate the impact of treatment outcomes on society via social validation procedures.

Clinical aphasiologists are, for the most part, novices in the arena of within-subject experimental designs, and we need to proceed cautiously in order to make maximal use of these designs to verify the effectiveness of our intervention procedures. We need to adopt a longitudinal perspective for our functional analysis research endeavors, and make a commitment to using the data from each experimental analysis to generate subsequent research questions. We need to be cooperative and collaborative in our investigations of aphasia, and view the role of replication as a scientifically appropriate and vital aspect of the experimental process. One-time efforts in within-subject analyses of aphasia intervention often yield interesting results, but if we want to make unequivocal statements about what our treatments can and will do, once is not enough.

REFERENCES

Davis, G.A. The clinical application of withdrawal, single-case research designs. In R.H. Brookshire (Ed.), Clinical Aphasiology: Conference Proceedings, 1978. Minneapolis: BRK Publishers, 1978.

Hersen, M. and Barlow, D.H. Single Case Experimental Designs: Strategies for Studying Behavior Change. New York: Pergamon Press Inc., 1976.

Johnston, J.M. and Pennypacker, H.S. Strategies and Tactics of Human Behavioral Research. Hillsdale, N.J.: Lawrence Erlbaum, 1980.

Kazdin, A.E. Single-Case Research Designs: Methods for Clinical and Applied Settings. New York: Oxford University Press, 1982.

Kearns, K. and Salmon, S. An experimental analysis of auxiliary and copula verb generalizations. J. Speech Hearing Dis., 1983 (in press).

McCarthy, P.A. Personal communication, 1983.

McReynolds, L.V. and Kearns, K.P. Single-Subject Experimental Designs in Communicative Disorders. Baltimore: University Park Press, 1983.

Post, J. and Leith, W. I'd rather tell a story than be one. Asha, 25, 23-26, 1983.

Sanders, S.B., Davis, G.A. and Hubler, V. A study of the interdependence of word retrieval and repetition in conduction aphasia. In R.H. Brookshire (Ed.), Clinical Aphasiology: Conference Proceedings, 1979. Minneapolis: BRK Publishers, 1979.

Sidman, M. Tactics of Scientific Research: Evaluating Experimental Data in Psychology. New York: Basic Books, 1960.

Stanton, K.M., Yorkston, K.M., Aune, K.J., and Hedges, J.S. Error recognition utilized to improve written language in a head injured patient. In R.H. Brookshire (Ed.), Clinical Aphasiology: Conference Proceedings, 1982. Minneapolis: BRK Publishers, 1982.

Thompson, C.K., McReynolds, L.V., and Vance, C.E. Generative use of locatives in multiword utterances in agrammatism: A matrix-training approach. In R.H. Brookshire (Ed.), Clinical Aphasiology: Conference Proceedings, 1982. Minneapolis: BRK Publishers, 1982.

Tonkovich, J.D. and Berman, M.S. Use of the two-alternative forced-choice paradigm in training aphasic error recognition. In R.H. Brookshire (Ed.), Clinical Aphasiology: Conference Proceedings, 1981. Minneapolis: BRK Publishers, 1981.

Tonkovich, J.D. and Loverso, F. A training matrix approach for gestural acquisition by the agrammatic patient. In R.H. Brookshire (Ed.), Clinical Aphasiology: Conference Proceedings, 1982. Minneapolis: BRK Publishers, 1982.

Wolf, M.M. Social validity: The case for subjective measurement or how applied behavioral analysis is finding its heart. J. Appl. Behav. Anal., 11, 203-214, 1978.