Single-Case Research Designs in Clinical Child Psychiatry

ALAN E. KAZDIN, PH.D.

Single-case designs refer to a methodological approach that can be used to investigate the efficacy of treatment with the individual patient. The designs permit scientifically valid inferences to be drawn about the effects of treatment and hence offer advantages over alternative strategies such as the uncontrolled case study or open study that are used with the individual case. The present article discusses the need for and utility and requirements of single-case designs in clinical child psychiatry. The underlying rationale and essential characteristics are presented and illustrated in several clinical cases. The feasibility of integrating assessment and design requirements of single-case research into clinical practice, the ethical issues such an integration may raise, and generality of findings from research on the single-case are also discussed.

Journal of the American Academy of Child Psychiatry, 22, 5:423–432, 1983.

The case study, or the intensive study of the individual patient, has played a major role in generating knowledge in several areas of psychiatry including the diagnosis, description, and treatment of various disorders. Modern psychiatric diagnosis began with careful analysis and the accumulation of individual cases (Kraepelin, 1883). For disorders that are relatively rare (e.g., multiple personality), information has come primarily from careful descriptions of individual cases (e.g., Thigpen and Cleckley (1957)). Perhaps the greatest impact of the case study has been in the area of treatment of clinical disorders. The history of many psychotherapy techniques or the conceptual models on which they are based can be traced to the influence of one or a few cases (e.g., Little Hans, Anna O. for psychoanalytic therapy; Little Albert, Peter for behavior therapy).

The case study serves as an important source of hypotheses about treatment and its effects. Yet, it is also recognized to be unacceptable as a method of obtaining scientific knowledge. Uncontrolled case studies characteristically rely on anecdotal information such as clinical impressions, judgments, and inferences. They lack the experimental control procedures that are required to draw scientifically valid inferences (Campbell and Stanley, 1963). For example, in the context of treatment, change in a patient's dysfunction may be marked. Yet, given the way in which treatment is implemented and evaluated, it may be impossible to rule out the influence of extraneous factors, other than treatment, that might account for change. The types of controls utilized in research usually are not possible, feasible, or permissible in clinical work with individual patients because of practical as well as ethical constraints.

Recently, a research methodology has emerged that is referred to as single-case experimentation. Although the methodology can be applied to large groups, it is uniquely applicable to the scientific evaluation of treatment for the individual patient. Use of the designs has been increasingly advocated in clinical and applied work, as attested to in part by recent publications in psychiatry and clinical psychology (Chassan, 1979; Hersen and Barlow, 1976; Kazdin, 1982), social work (Jayaratne and Levy, 1979), and education and counselling (Kratochwill, 1978). The purpose of the present paper is to discuss single-case research as a methodology to evaluate treatment and to illustrate its applicability to clinical work in child psychiatry. The paper discusses the underlying rationale and characteristics of the methodology and its compatibility with several features of clinical work.

Need for and Utility of Single-Case Designs

Research on the treatment of clinical disorders usually is conducted in academic or academically affiliated settings (Parloff, 1979). Yet, the bulk of treatment is practiced in hospitals and clinics. The different priorities, resources, and clinical exigencies obviously make research more feasible in academic rather than clinic settings. The problem for developing the scientific basis for clinical treatment is that much of the rigorous work is conducted under conditions that depart from the usual clinical practice. In research settings, several conditions often are invoked

Dr. Kazdin is Professor of Child Psychiatry and Research Director of the Child Psychiatric Treatment Service at the Western Psychiatric Institute and Clinic, University of Pittsburgh School of Medicine (3811 O'Hara St., Pittsburgh, PA 15213), where reprints may be requested.

Completion of this paper was supported by a Research Scientist Development Award (MH00353) and a Grant (MH35408) from the National Institute of Mental Health.

0002-7138/83/2205-0423 $02.00/0

to meet the requirements of between-group experimental designs such as identifying homogeneous groups of patients, assigning patients randomly to treatments, providing a standardized treatment regimen, as specified by the protocol, and so on (APA Commission on Psychotherapies, 1982). The patients who participate often are carefully screened so that they present a relatively homogeneous diagnostic picture to which a treatment protocol is aimed. When the investigation is completed, the effects of treatment usually are evaluated by comparing alternative groups statistically to see if mean (average) patient performance varies between (or among) groups.

The situation confronting the clinician departs considerably from the circumstances available to the researcher. The practicing clinician is confronted with the individual case (e.g., one child, family). Conclusions derived from group research about the average response of patients to a particular form of treatment are difficult to apply to the individual. Also, the clinician cannot apply a particular treatment in a standardized fashion to a circumscribed problem. Patients may present multiple problem areas that require a flexible, multifaceted, and individually tailored intervention.

If the clinician were interested in evaluating the effects of treatment, there has been no scientifically acceptable methodology that is congruent with treatment demands of the individual case. The usual means of reporting information is the case study, where anecdotal information is presented in a descriptive fashion. Although case studies can generate valuable hypotheses for research, they lack sufficient control to identify the bases for therapeutic change. Even if several cases are accumulated in a single report and assessment is introduced to examine clinical changes before and after treatment—as is sometimes the case in an open study—interpretation of the results remains ambiguous.

The unique feature of scientific research is to make implausible or to rule out alternative explanations that might be proffered to account for a particular relationship. In treatment research, the purpose is to rule out extraneous factors that explain changes the clinical investigator wishes to attribute to treatment. Patient improvement over time might result from several nontreatment influences such as events in the child's life (e.g., increase or decrease in stress at home or at school), or events within the individual (e.g., maturation, remission of symptoms), and a variety of other methodological and statistical artifacts commonly discussed in research (see Cook and Campbell (1979) and Kazdin (1980)). Research designs that compare groups of patients who receive alternative treatment and control (e.g., no treatment, placebo) conditions make implausible the impact of such extraneous factors through such practices as random assignment of patients to groups and assessment of performance before and after treatment.

For the practicing clinician, there has been no scientific research methodology that can be used with individual patients. Single-case designs represent a scientific methodology that can evaluate alternative treatments and rule out the impact of extraneous factors as rival explanations of the results. More importantly, the methodology provides a flexible approach to evaluation that is consistent with many of the priorities, professional responsibilities and practical exigencies of clinical practice.

Characteristics and Requirements of Single-Case Designs

Single-case designs permit the evaluation of treatments as they are applied clinically. Inferences are drawn by utilizing the patient as his or her own control. This means that the impact of treatment is examined in relation to the patient's symptoms or dysfunction over time. There are, of course, special requirements that need to be invoked to permit evaluation. But the evaluation can be designed and implemented in a variety of ways to meet the demands of the individual case. The treatment focus, means of evaluating progress, specific intervention technique(s), and how and when treatment is applied can be arranged in such a way as to retain the priority of service delivery.

Specification of the Treatment Focus

A major requirement of single-case designs is specification of the goals of treatment or those symptoms and areas of functioning that are to be altered. Operationally, this means that measures need to be selected that would be expected to reflect progress in treatment. There are no inherent limitations regarding the facets of the patient's symptoms that are assessed in single-case research. Measures of cognition, affect, behavior, psychophysiology, or personality dimensions can be used. Yet, the clinician needs to identify prior to applying treatment what sorts of changes will be used to evaluate progress. For example, for a depressed child, alternative self-report and interview measures might be administered to the child and/or his or her parents (Kazdin and Petti, 1982). A self-report measure or interview alone or in combination with other measures (e.g., side effects checklist) might be used to evaluate response to treatment. Practical constraints may dictate the feasibility of using certain measures. In outpatient treatment, self-report, parent or teacher ratings, might be particularly feasible. An important determinant of selection of the measure is

whether it can be administered or obtained on several occasions, as discussed below.

Continuous Assessment

The most fundamental design requirement of single-case experiments is the use of repeated observations of performance over time. The measures selected to evaluate progress in treatment are administered on several occasions, usually before the treatment is applied and continuously over the course of treatment. Ideally, assessment is conducted on a daily basis or at least on multiple occasions each week.

Continuous assessment is a basic requirement because single-case designs examine the effects of interventions on performance over time. Continuous assessment allows the clinician to examine the pattern and stability of performance before treatment is initiated. The pretreatment or baseline information provides a picture of what performance is like without the treatment. The observations are continued so that the clinician can examine changes on the measures as treatment is applied.

The role of continuous assessment in single-case research can be conveyed by examining a basic difference between-group and single-case research. In group research, one group might receive treatment while a control group does not. The question of whether treatment produces change is evaluated by collecting one or two observations (pre- and post-treatment assessment) for all persons in each group. In single-case research, the effects of the treatment can be examined by observing the influence of treatment and no treatment on the performance of the same person. Instead of one or two observations of several persons, as in group research, several observations are obtained for one or a few persons. Continuous assessment over time provides the observations required to compare performance under different conditions (e.g., treatment and no treatment or alternative treatment and placebo conditions). The logic of between-group and single-case designs does not differ fundamentally. Both design strategies require making comparisons of performance under different conditions. Continuous assessment over time, as used in single-case designs, permits the comparisons to be made within the subject, i.e., using the patient as his or her own control.

Separate Phases

In single-case designs, assessment is continued over the course of different phases, or periods in which different conditions are in effect. Typically, the designs begin with a period of assessment referred to as the baseline (pretreatment) phase. Baseline data are obtained on several occasions before treatment is initiated. The baseline serves different functions in single-case research. First, baseline data describe the existing level of the problem and hence serve a *descriptive function.* Second, the data serve as a basis for predicting the likely level of performance in the immediate future if treatment is not provided. This *predictive function* is critical to the logic of single-case designs and warrants additional comment.

To evaluate the impact of treatment in single-case research, it is important to have an idea of what the patient's performance would be like in the immediate future without the intervention. An extrapolation of the baseline level of performance suggests the likely course of the symptoms in the immediate future. Several days of observation are usually needed to identify stable performance on the assessment device for this projection to be made. On most measures, performance is likely to fluctuate unsystematically. Also, the possibility exists that symptoms may be improving or becoming worse over time. Baseline data provide a measure of the level of the patient's performance and the direction of change, if any, that the symptoms may show in the short run.

After the baseline phase is completed, treatment is implemented. Data in the treatment phase also serve separate functions. Performance on the measures during the treatment phase is compared with the projected level of performance from the baseline phase. If the patient's performance departs markedly from what whould be expected from extrapolation of the baseline level of performance, this is consistent with the view that treatment has produced an effect. Essentially, the projected level of baseline serves as a criterion to evaluate if treatment has led to change. Presumably, if treatment is effective, performance during the treatment phase will differ from the projected level of baseline.

The design usually continues beyond the mere assessment of performance under separate baseline and treatment phases. Changes in performance across baseline and treatment phases might be due to a change in the patient's life or environment that is unrelated to the application of treatment. The purpose of single-case research and indeed all research is to rule out alternative explanations of the data. In single-case research, the situation is arranged so that the influence of extraneous factors that might explain the changes from baseline to treatment phases are ruled out or made implausible. These arrangements consist of the different experimental designs.

Stability of Performance

Since baseline performance is used to predict how the patient will respond in the immediate future, it is important that the data are relatively stable. A stable rate of performance means the absence of a trend and

relatively little variability (fluctuation). A trend refers to a systematic increase or decrease in performance over time. For example, the gross motor activity of an attention deficit disorder child may systematically decline during baseline. A trend toward improvement during baseline may make evaluation difficult, because it is in the same direction of the anticipated treatment effects. A veridical treatment effect might be difficult to distinguish from the projected level of baseline data. Thus, in single-case research, baseline data ideally show little or no trend or a trend in the opposite direction from what would expected during treatment.

Stability of performance also refers to fluctuation or variability in the patient's performance over time. Excessive variability during baseline or other phases may interfere with drawing conclusions about treatment effects. As a general rule, the greater the variability in the data, the more difficult it is to draw conclusions about the effects of the treatment.

Excessive variability is a relative notion. Whether the variability is excessive and interferes with drawing conclusions about treatment depends on many factors such as the initial level of performance during baseline and the extent of change during treatment. In the extreme case, baseline performance may fluctuate daily from extreme high to low levels (e.g., 0–100% on a self-report inventory or observed behavior at home or in a classroom). With such extreme fluctuations in performance, it may be difficult to predict a stable level of performance in the immediate future. Of course, the extreme fluctuations may be clinically significant in their own right and be altered with treatment.

Obtaining stable performance data during baseline and treatment phases usually is not a problem. The designs can handle a considerable degree of instability because trends in the data and fluctuation in performance themselves usually change in response to different interventions or phase changes in single-case designs. So, even when there is a baseline trend in the direction of therapeutic change or a high degree of fluctuation in performance, single-case designs can still be applied. In special instances where data may be difficult to evaluate, special statistical techniques may be used to clarify the data (see Kazdin (1982)).

Illustrations

The essential characteristics of single-case designs and how they can be used to draw inferences about treatment in clinical work are conveyed by illustrations in both inpatient and outpatient applications. Two commonly used single-case designs are illustrated to convey the logic and application of single-case research methods, but they do not exhaust the available options.

ABAB or Reversal Designs

One commonly used design strategy is referred to as an ABAB or reversal design. Separate phases (A = baseline or no intervention, B = treatment) are alternated over time to demonstrate if treatment is responsible for change. For example, Zlutnick et al. (1975) evaluated a procedure to reduce seizures in several children, each evaluated individually in separate ABAB designs. One case was a 7-year-old boy with a 5-year history of seizures and diagnosis of autism and brain damage. Seizures averaged 12 per day despite anticonvulsant medication (Dilantin®). To control seizures, an interruption procedure was used whenever preseizure activity (fixed stare, body rigidity, violent shaking) occurred. The procedure, carried out in a special classroom at school, consisted of saying "no" loudly to the child and holding his shoulders with both hands and shaking him once when the stare began. This procedure was designed to interrupt the sequence leading to a seizure. The interruption procedure was evaluated in an ABAB design.

After baseline assessment of seizures in class, the interruption procedure was implemented. The effects of the procedure can be seen by comparing seizures during the baseline and treatment phases (see fig. 1). To evaluate if the treatment, rather than extraneous factors coincident with the onset of treatment, accounted for the changes, the procedure was temporarily suspended. This is referred to as a reversal or return-to-baseline condition. When treatment was suspended, seizures returned to near baseline levels. The interruption procedure was reinstated until seizures were eliminated. Over the course of a 6-month follow-up, after the procedure had been terminated, only one seizure had occurred. Overall, the effects of the intervention were clearly demonstrated. The results indicated dramatic changes whenever the interruption procedure was introduced. The pattern of the data makes implausible the effects of extraneous factors (e.g., spontaneous improvements over time) in explaining the results.

The crucial feature of an ABAB design is the alternation of phases over time to assess whether treatment produces change. Several different phases and treatments may be applied to achieve the clinical goals. For example, Wells et al. (1981) evaluated the effects of alternative procedures alone and in combination to alter the behavior of a hyperactive 9-year-old boy hospitalized on an inpatient service. The basis of referral included poor concentration and unmanageable behavior at school and aggression toward peers and adults. In an ABAB design, alternative medications (Dexedrine® and Ritalin®) were evaluated on classroom performance in the hospital. In addition to medication, treatment eventually included a self-con-

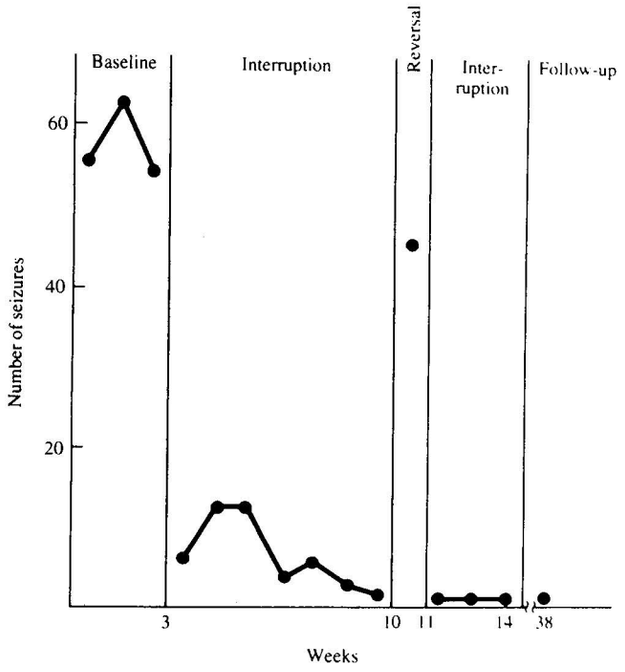

FIG. 1. The number of seizures per week in separate phases of an ABAB design (A = baseline, B = interruption procedure). Follow-up data reflect the total number of seizures for the 6-month period following termination of treatment. (Reproduced with permission from S. Zlutnick et al.: *Journal of Applied Behavior Analysis*, 8:1–12, 1975.)

trol regimen in which the child was trained to monitor his own on-task behavior. Whenever a tone sounded in class, the child was told to take a token (poker chip) from a bowl near him if he was on task. The effects of medication and self-control treatments were evaluated on several behaviors in class (see fig. 2). The results indicated that medication alone (Ritalin®) produced some changes. When the self-control procedure was combined with medication, marked changes occurred (see CD phase). To evaluate if the combination of medication and self-control procedures were necessary, placebo and self-control procedures were implemented which led to an increase in hyperactive behaviors. In the final phase, medication and self-control were reinstituted and marked changes in behavior were evident. This case demonstrated treatment effects each time the combined procedure was introduced. Importantly, continuous assessment of performance over time revealed that an additional intervention was needed with medication to produce clinical improvement.

The ABAB design and its variations nicely convey the logic of single-case designs, because they show how performance during baseline and treatment phases is compared. If performance changes in response to alternation of baseline and treatment phases, the most plausible explanation is that the intervention rather than extraneous influences produced the change. However, the designs often are not clinically practical. Showing that performance reverts to baseline levels when treatment is temporarily withdrawn is tantamount to making the child worse. Alternative single-case designs avoid this problem.

Multiple-Baseline Designs

A frequently used design strategy, referred to as a multiple-baseline design, demonstrates the effect of treatment without reverting to baseline conditions. The unique feature of multiple-baseline designs is the introduction of treatment sequentially across different facets or performance. Essentially, two or more baselines are assessed (e.g., symptom areas or performance across different situations). The effect of treatment is demonstrated if changes are associated with the application of treatment to the specific symptoms or situation as treatment is extended in a sequential fashion.

For example, Kandel et al. (1977) treated a severely withdrawn boy (diagnosed as autistic) enrolled in a school for children with speech, hearing, and emotional disorders. At school, the child spent his free time alone, often talking to himself. Treatment was designed to increase social interaction with his peers. Treatment was introduced into two free-play periods at school. Treatment consisted of first having a trainer model or demonstrate appropriate social interaction for the child. Two other children were included and were encouraged to play with the boy. Rewards (candy) were given to the two children for their help with training. To demonstrate that the intervention was responsible for change, it was introduced to the two free-play periods at school at different points in time (on the playground and juice time). As shown in figure 3, social interaction increased in each of the situations when treatment was introduced. The increases in interaction in each situation only when training was introduced strongly suggests that the training program was responsibe for change. Follow-up evaluation, conducted 3 weeks later when the program was no longer in effect, showed that the changes were maintained. A 9-month follow-up (upper portion of fig. 3) was obtained after the child had been attending a regular school where free time was observed. The high levels of social interaction were maintained at the regular school.

As a final illustration, Frame et al. (1982) evaluated treatment with a 10-year-old boy with a diagnosis of major depression (DSM-III) and hospitalized on an inpatient unit. Interviews with the child indicated that he avoided social interaction with others, as evident in poor eye contact, bodily positions that turned away from others, inaudible and constricted speech, and

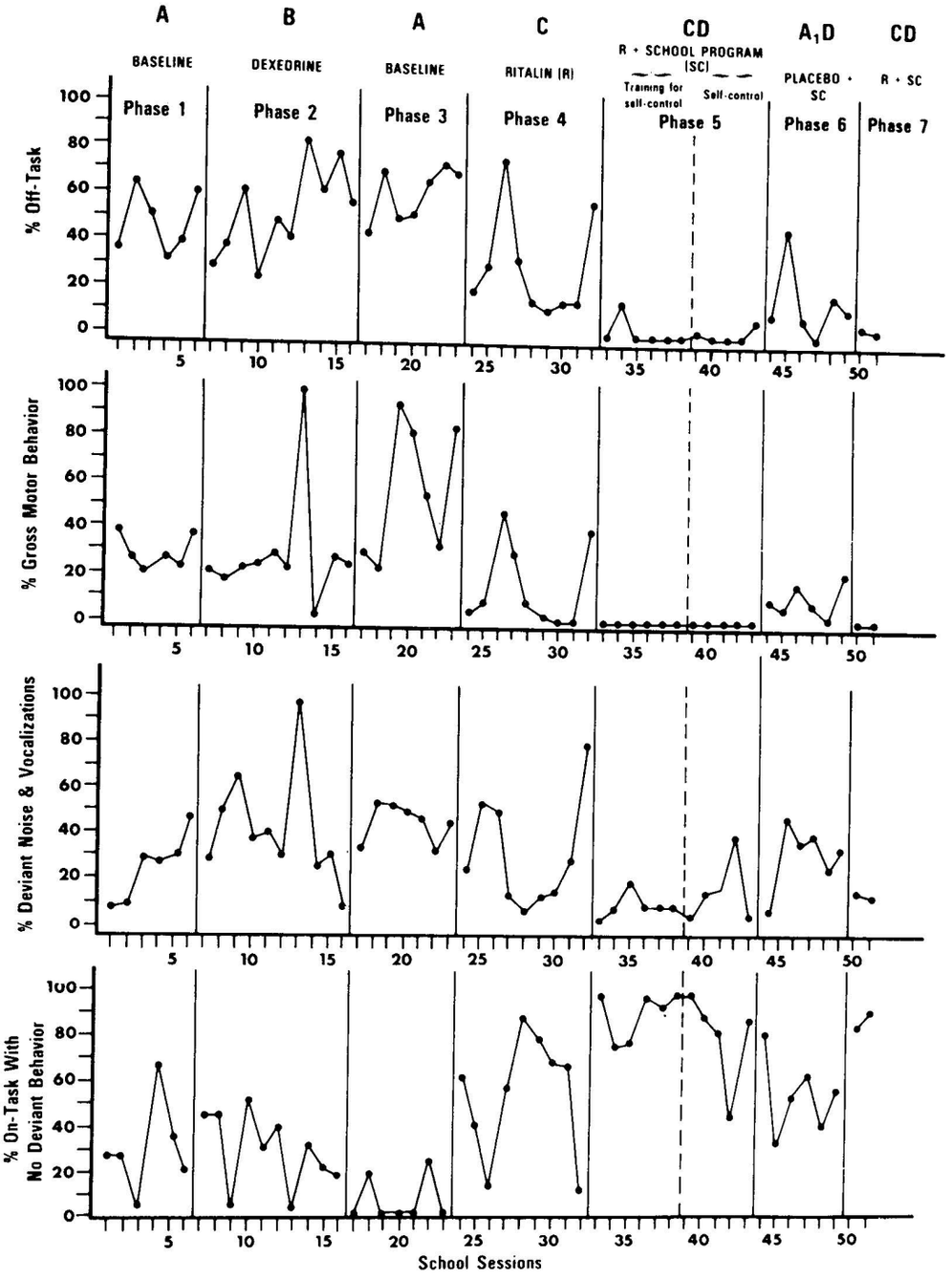

FIG. 2. Percent occurrence in the classroom of off-task, gross motor behavior, deviant noise and vocalizations, and on-task behavior with no other deviant behavior recorded. The data were gathered in a variation of an ABAB design with baseline (A phase) and alternative treatment (B, C, D alone or in combination) phases. (Reproduced with permission from K. C. Wells et al.: *Behavioral Assessment, 3*:359–369, 1981.)

bland affect in his verbalizations. There areas served as the focus of treatment which was provided in individual sessions 5 times per week for approximately 5 weeks. Assessment interviews were continued throughout the treatment at which point the specific social behaviors were evaluated by blind raters. Treatment conducted by a therapist consisted of training the child to interact in a variety of interpersonal situations. The therapist modeled the desired performance, allowed the child the opportunity to practice, and provided praise and feedback as needed. The treatment was introduced in a multiple-baseline

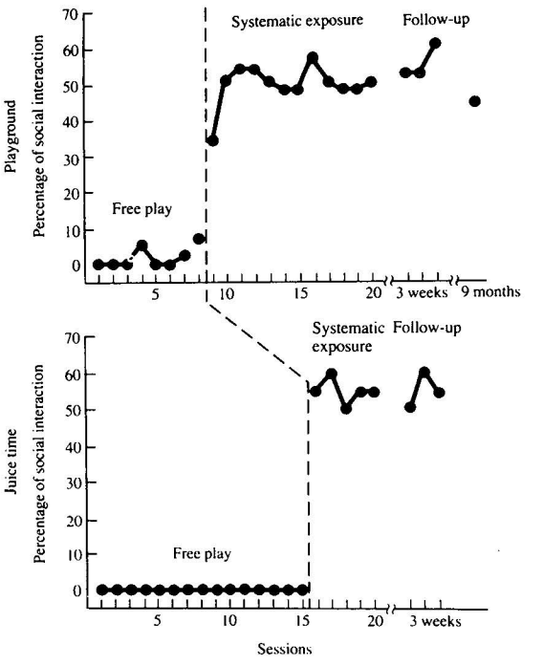

FIG. 3. Social interaction at school in two separate free-play situations, on the playground and in the courtyard at juice time. The intervention was introduced to each situation at different points in time to meet the requirements of multiple-baseline design across situations. (Reproduced with permission from H. J. Kandel et al.: *Journal of Behavior Therapy and Experimental Psychiatry, 8*:75–81, 1977.)

across behaviors in which the separate behaviors were focused on at different points in time.

As shown in figure 4, when training was introduced to the first two behaviors (eye contact and body position), marked changes occurred. When training focused on other facets of interaction, these changed as well. Follow-up evaluation 12 weeks after treatment had been terminated indicated that the gains were maintained. In general, the pattern of change suggested that the intervention, rather than extraneous factors such as contact with the therapist or repeated exposure to social situations in training, accounted for the change.

General Comments

The preceding illustrations convey only two of the many different design options available in single-case research (see Hersen and Barlow (1976) and Kazdin (1982)). Each design is characterized by continuous assessment and replication of intervention effects over time. Also, each design compares performance under treatment or no-treatment or under alternative treatment conditions. All research involves a comparison of some sort. In ABAB designs, the comparison usually involves performance under baseline or reversal (no treatment) and treatment conditions. Predictions are made about what performance would be like without the intervention by extrapolation of baseline performance. Changes evident with the presentation, withdrawal and re-presentation of treatment test whether the intervention is likely to be responsible for change. In multiple-baseline designs, the untreated symptoms serve as no treatment control conditions to evaluate the changes that could be expected without the application of treatment. The repeated demonstration of change as the intervention is applied consecutively to different symptoms or in different situations increases the plausibility that the intervention rather than extraneous factors was responsible for change.

Issues and Limitations

Clinical Feasibility

The purpose of the above illustrations was to convey the use of single-designs. The specific examples depended on careful evaluation and assessment resources that might not be feasible in many routine outpatient applications of treatment. Hence, the clinician may take the examples as further proof that careful evaluation and research in clinical work remain impractical. In many cases the clinician may only be able to approximate features of single-case designs because of difficulties in implementing various

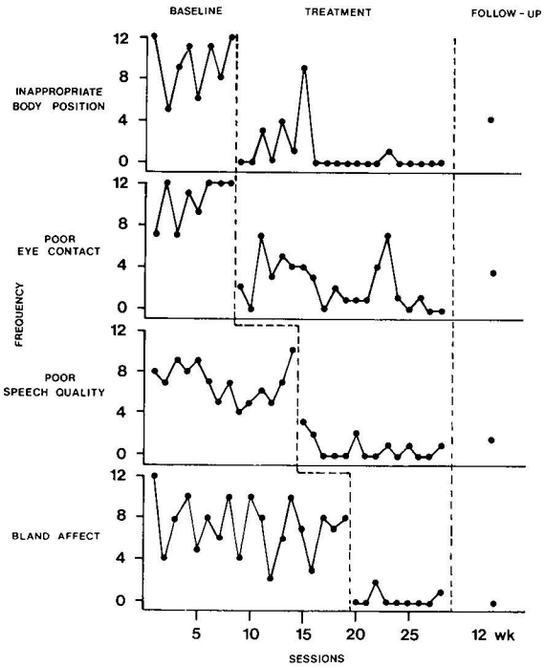

FIG. 4. Frequency of specific features of social behaviors of a depressed boy as measured during assessment interviews. The treatment was applied to different behaviors at different points in time to meet the requirements of a multiple-baseline design across behaviors. (Reproduced with permission from C. Frame et al.: *Journal of Behavior Therapy and Experimental Psychiatry, 3*:239–243, 1982).

assessment strategies or collecting information on a daily basis. Yet, compromises can be reached that still retain critical features of the methodology within particular constraints that the case or treatment situation presents (Hayes, 1981; Kazdin, 1981).

For example, daily assessment of the child's symptoms may not be possible in outpatient treatment. Yet, it may be feasible for a teacher or parent to complete a brief rating form to evaluate the child's performance. The teacher or parent could repeatedly complete portions of standardized checklists (e.g., Achenbach (1978) and Conners (1969)) or a diagnostic interview (e.g., Chambers et al. (1978)) that evaluates the child's symptoms as they occurred within a 1- or 2-day period. Also, when the child is seen in treatment, he or she can complete a self-report questionnaire or interview to assess particular symptoms on a weekly basis. The purpose of assessment is to determine the pattern of performance over time and whether a particular form of treatment affects that pattern. The purpose can be achieved even if daily assessment is not feasible.

Use of particular single-case designs may not be feasible with a given clinical case. For example, it is not usually feasible to apply treatment in an ABAB design where treatment is temporarily withdrawn after improvements have been achieved. A multiple-baseline design may not be possible either if the treatment is not expected to produce highly specific effects across only select symptoms or situations. Yet, features of alternative designs can be used to improve the inferences one can draw from treatment. For example, baseline observations can be followed by separate phases where variations of treatment (e.g., different medication doses) or alternative treatments (e.g., medication and psychotherapy) are combined until sufficient clinical improvement is achieved. Using the patient as his or her own control and obtaining continuous measures allow the clinician to examine the comparative effects of alternative variations of treatment.

Discussion of the feasibility of using single-case designs in clinical work has another side. Features of single-case designs are not only feasible but offer important benefits for clinical work. Continuous assessment provides ongoing data about treatment progress and yields critical information to the clinician. The clinician can evaluate if treatment is having its intended effects. And, if treatment is producing change, the clinician can evaluate if the change is sufficient to be clinically important. If not, changes can be made in treatment or alternative treatments can be applied, as illustrated in one of the above examples (Wells et al., 1981).

Single-case designs also permit gradual introduction of treatments to change specific symptom areas or performance in select situations, as illustrated in multiple-baseline designs. Essentially, applications of treatment in a multiple-baseline design provides a preview of treatment effects implemented for a circumscribed aspect of performance. Treatment can be extended to other areas of performance or situations only if initial information (on the first baseline) suggests it is producing change.

Generality of the Results

A major concern with single-case research is that the results may not be generalizable to persons other than the one included in a particular clinical trial. The generality of single-case research is often discussed in relation to between-group research. Because between-group research uses larger numbers of subjects than does single-case research, the findings are often assumed to be more generalizable. Yet, the use of a larger number of subjects does not, by itself, ensure generalizable findings (Kazdin, 1982; Sidman, 1960). In the majority of between-group investigations, results are evaluated on the basis of the average group performance. Group analyses do not shed light on the generality of treatment effects among patients within a group.

For example, if 20 patients who receive treatment show an average change greater than 20 other patients who receive no treatment, little information is available about the generality of the results. We do not know by this group analysis alone how many persons in the treatment group were affected or affected in a clinically important way. Ambiguity about the generality of findings from between-group research is not inherent in this research approach. However, investigators infrequently examine individual patient data as well as group data to make inferences about the generality of effects among patients within a given treatment condition. If the individual data were examined in between-group research, a great deal might be said about the generality of the findings.

It is quite possible that the effects of treatment demonstrated within the individual patient may not generalize to other patients. Indeed, this is not a unique problem or possibility raised by single-case research but evident in clinical applications that focus on one or a few cases. The traditional uncontrolled case study of course reflects the same potential problem, namely, that the findings might be restricted uniquely to that case. The benefit of the single-case methodology is that it permits much more to be said about the effects of treatment for that case.

Although individuals respond differently to a par-

ticular treatment regimen, the efficacy of alternative treatments is not entirely idiosyncratic. Generality of treatment effects, no doubt, varies as a function of the strength, dose, duration, and other aspects of treatment as well as characteristics of the patients and disorders to which treatments are applied. The potential problem of limited generality of results in single-case research has not in fact been found in areas where the methodology has been applied. For example, single-case designs have been frequently used in evaluating selected behavior therapy techniques. Treatments demonstrated to be effective in single-case applications often have been shown to have wide generality among clinical populations, settings, and problems (e.g., Kazdin (1982) and Kazdin and Tuma (1982)).

The ultimate test of generality of findings both in clinical work and research is replication, or repeated demonstration of the effects of treatment. Replication examines the extent to which results obtained in one demonstration extend (or can be generalized) to other cases, populations, settings, measures, disorders, clinicians, and treatment sites. The way to answer questions of generality of single-case research is the same as it is for group research, namely independent replications of intervention effects. Clinicians in their own practice can accumulate cases using single-case designs and can directly test the generality of particular treatment strategies.

Ethical Issues

Research, whether based on single-case or group designs, can place restrictions on the administration of treatment in such a way as to raise important ethical issues. For example, in group research withholding treatment (as part of a no-treatment or waiting-list control group) is ethically undesirable if patients are in need of immediate treatment. Different facets of single-case designs might also seem to compete with the ethical and professional responsibilities of service delivery. For example, when a patient is in need of treatment, it may not be feasible to conduct baseline assessment for an extended period (e.g., 2 weeks). The urgency of clinical intervention may require immediate treatment (e.g., medication for a depressed and suicidal adolescent). However, obtaining baseline information before applying treatment is a matter of degree. Not all clinical problems may require immediate clinical intervention. Obtaining baseline information, even if only for a brief period, can provide information that is useful for evaluating treatment progress. Indeed, baseline assessment, rather than being an obstacle to clinical treatment, might be more appropriately viewed as part of a thorough clinical workup. Information about the problem over time provides an important comparative base for evaluating subsequent performance and treatment efficacy.

Continuous assessment before and during treatment allows the clinician an objective and operational basis for evaluating ongoing treatment and making decisions about how to proceed further. Treatment can be stopped, continued, or altered as a function of the assessment information. An ethical issue might be raised by not evaluating treatment through some means of continuous assessment when such assessment is possible. Without careful evaluation, the possibility exists that the patient is not receiving an effective or maximally effective treatment. Assessment provides information to aid the clinician in evaluating whether treatment in fact is achieving its intended effects and serves a direct benefit to both clinician and patient alike.

Conclusions

Single-case designs provide a viable research methodology that in varying degrees can be incorporated into clinical practice. Key requirements for use of the designs are evaluation of patient progress over time and comparison of patient performance before and during treatment. These features are implicit in clinical practice. Single-case designs systematize the manner in which these features are incorporated into practice. Systematically evaluating patient dysfunction over time and applying treatment according to the design criteria reduce the plausibility that extraneous factors other than treatment could account for change. Consequently, the strength of the inferences that can be drawn is much greater than would be available from the usual uncontrolled case study. Apart from its research purposes, the data from single-case designs provide useful information to evaluate whether therapeutic changes have been achieved and whether the magnitude of these changes is sufficient to constitute a clinically significant change.

It is sometimes assumed that clinicians generally are not interested in evaluating their treatments. The lack of rigorous evaluation of most treatments as they are applied clinically does not necessarily support this assumption. In fact, a research methodology has not been available for evaluating applications of treatment with the individual case. The requirements of group research such as the accumulation of homogeneous populations to which standard treatments can be applied and the resources for a large-scale protocol are simply unavailable in clinical practice. For evaluation to be possible, a methodology is needed that is consistent with characteristics, demands and priorities of the clinical situation.

Single-case designs provide such a methodology. The design strategies serve as an intermediate step between uncontrolled case or open studies and large-scale clinical trials. Single-case designs do not replace between-group trials of alternative treatments. Indeed, single-case demonstrations and larger group studies represent complementary strategies for clinical research. They can also function reciprocally. Single-case demonstrations can identify causal relations between innovative treatments and therapeutic change and suggest treatments that warrant larger-scale investigation. Alternatively, evidence obtained from large-scale trials can be tested clinically with the single-case to evaluate the impact of treatment on individual patients. Single-case designs uniquely provide the opportunity to evaluate treatment in clinical practice where the most demanding conditions of treatment exist. The methodology when used in clinical work is not only likely to build a knowledge base but also can provide the opportunity to test highly innovative hypotheses that emerge from the richness of clinical experience.

References

ACHENBACH, T. M. (1978), The child behavior profile: I. Boys aged 6–11. *J. Consult. Clin. Psychol.*, 46:478–488.

AMERICAN PSYCHIATRIC ASSOCIATION COMMISSSION ON PSYCHOTHERAPIES (1982), *Psychotherapy Research: Methodological and Efficacy Issues.* Washington, D.C.: American Psychiatric Association.

CAMPBELL, D. T. & STANLEY, J. C. (1963), *Experimental and Quasi-Experimental Designs for Research.* Chicago: Rand McNally.

CHAMBERS, W., PUIG-ANTICH, J. & TABRIZI, M. A. (1978), The ongoing development of the Kiddie-SADS (Schedule for Affective Disorders and Schizophrenia for School-age Children). Presented at the American Academy of Child Psychiatry, San Diego, Calif.

CHASSAN, J. B. (1979), *Research Design in Clinical Psychology and Psychiatry*, ed. 2. New York: Irvington.

CONNERS, C. K. (1969), A teacher rating scale for use in drug studies with children. *Amer. J. Psychiat.*, 126:884–888.

COOK, T. D. & CAMPBELL, D. T. (eds.) (1979), *Quasi-Experimentation: Design and Analysis Issues for Field Settings.* Chicago: Rand McNally.

FRAME, C., MATSON, J. L., SONIS, W. A., FIALKOV, M. J. & KAZDIN, A. E. (1982), Behavioral treatment of depression in a prepubertal child. *J. Behav. Therapy Exp. Psychiat.*, 3:239–243.

HAYES, S. C. (1981), Single-case experimental design and empirical clinical practice. *J. Consult. Clin. Psychol.*, 49:193–211.

HERSEN, M. & BARLOW, D. H. (1976), *Single-Case Experimental Designs: Strategies for Studying Behavior Change.* New York: Pergamon.

JAYARATNE, S. & LEVY, R. L. (1979), *Empirical Clinical Practice.* New York: Columbia University Press.

KANDEL, H. J., AYLLON, T. & ROSENBAUM, M. S. (1977), Flooding or systematic exposure in the treatment of extreme social withdrawal in children. *J. Behav. Therapy Exp. Psychiat.*, 8:75–81.

KAZDIN, A. E. (1980), *Research Design in Clinical Psychology.* New York: Harper & Row.

—— (1981), Drawing valid inferences from case studies. *J. Consult. Clin. Psychol.*, 49:183–192.

—— (1982), *Single-Case Research Designs: Methods for Clinical and Applied Settings.* New York: Oxford University Press.

—— & PETTI, T. A. (1982), Self-report and interview measures of childhood and adolescent depression. *J. Child Psychol. Psychiat.*, 23:437–457.

—— & TUMA, A. H. (eds.) (1982), *New Directions for Methodology of Social and Behavioral Sciences: Single-Case Research Designs.* San Francisco: Jossey-Bass.

KRAEPELIN, E. (1883), *Psychiatrie.* Leipzig: Abel.

KRATOCHWILL, T. R. (ed.) (1978), *Single-Subject Research: Strategies for Evaluating Change.* New York: Academic Press.

PARLOFF, M. B. (1979), Can psychotherapy research guide the policymaker? A little knowledge may be a dangerous thing. *Amer. Psychol.*, 34:296–306.

SIDMAN, M. (1960), *Tactics of Scientific Research.* New York: Basic Books.

THIGPEN, C. H. & CLECKLEY, H. M. (1957), *Three Faces of Eve.* New York: McGraw-Hill.

WELLS, K. C., CONNERS, C. K., IMBER, L. & DELAMATER, A. (1981), Use of single-subject methodology in clinical decision-making with a hyperactive child in the psychiatric inpatient unit. *Behav. Assess.*, 3:359–369.

ZLUTNICK, S., MAYVILLE, W. J. & MOFFAT, S. (1975), Modification of seizure disorders: the interruption of behavioral chains. *J. Appl. Behav. Analy.*, 8:1–12.