American Journal of Clinical Hypnosis

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/ujhy20

A Methodological Review of Case Studies Published in the American Journal of Clinical Hypnosis

William R. Nugent M.S.W. a
a Florida State University, USA


To link to this article: http://dx.doi.org/10.1080/00029157.1985.10402607

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
A Methodological Review of Case Studies Published in the *American Journal of Clinical Hypnosis*

William R. Nugent  
*Florida State University*

A methodological framework for evaluating the internal and external validity of case studies is presented. Based on an evaluation of 74 case studies published in the *American Journal of Clinical Hypnosis* using this framework, the author makes recommendations for upgrading the value of case studies used and reported by practitioners using hypnotic interventions.

Case studies frequently appear in the clinical literature. Kazdin (1982, p. 88) defines the case study as "... uncontrolled reports in which one individual and his or her treatment are carefully reported and inferences are drawn about the basis of therapeutic change." Not only has the case study been used to report successes to colleagues (Hersen & Barlow, 1976), it has also been influential in the development of specific therapeutic techniques (Kazdin, 1980). The case study has been used as a report form, a tool in the development of clinical technique, and as a means of drawing causal inferences.

With the development of single case experimental designs, the case study has also come to refer to a methodological approach whereby an individual client (or group) and a relevant therapeutic intervention are studied in a manner precluding unambiguous conclusions concerning change and existence of any functional relationship between the intervention and observed effects (Campbell & Stanley, 1963; Paul, 1969; Kazdin, 1982). Generally, case studies use anecdotal information on client behavior and/or problems, the intervention used, and therapeutic change and are not structured to rule out other plausible (and sometimes contradictory) explanations for what was observed (Kazdin, 1982). Thus, conclusions about the occurrence of change and any causal inferences connecting treatment and change reflect merely the therapist’s personal opinion (Kazdin, 1982). Further, the anecdotal case study can limit efforts to replicate the intervention used and its generalization to other clients, problems, and therapists (Hersen & Barlow, 1976; Bloom & Fischer, 1982).

The author reviewed articles published in the *American Journal of Clinical Hypnosis* between January, 1973, and January, 1983. In total, 361 articles were published. Out of this total, 140 used case studies (38.8%). Over a five year span (1978 thru 1982), 74 of 175 articles used case studies (42.3%). A year-by-year tabulation can be seen in Table 1. As the data show, the case study is a report form and methodological approach frequently used by clinicians.
TABLE 1

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of Articles Involving Case Studies</th>
<th>Total Number of Journal Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>1973</td>
<td>13</td>
<td>33</td>
</tr>
<tr>
<td>1974</td>
<td>14</td>
<td>39</td>
</tr>
<tr>
<td>1975</td>
<td>14</td>
<td>38</td>
</tr>
<tr>
<td>1976</td>
<td>12</td>
<td>38</td>
</tr>
<tr>
<td>1977</td>
<td>13</td>
<td>38</td>
</tr>
<tr>
<td>1978</td>
<td>10</td>
<td>34</td>
</tr>
<tr>
<td>1979</td>
<td>14</td>
<td>32</td>
</tr>
<tr>
<td>1980</td>
<td>11*</td>
<td>26*</td>
</tr>
<tr>
<td>1981</td>
<td>22</td>
<td>44</td>
</tr>
<tr>
<td>1982</td>
<td>17</td>
<td>39</td>
</tr>
<tr>
<td>10 Year totals</td>
<td>140</td>
<td>361</td>
</tr>
<tr>
<td>5 Year totals</td>
<td>66</td>
<td>186</td>
</tr>
<tr>
<td>1973–1977</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Year totals</td>
<td>74</td>
<td>175</td>
</tr>
<tr>
<td>1978–1982</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Data missing for April 1980 Journal

using hypnotic interventions. No articles were found in which case study methodology was reviewed.

There are inherent weaknesses in the case study, as suggested above. Thus, it behooves practitioners using this methodology to evaluate the case study, make appropriate alterations in practice and report forms, and thereby increase the value of published clinical efforts for other professionals. In this article, the author develops a framework for evaluating the internal validity and replicability of case studies. Within this framework, 74 case studies published in the American Journal of Clinical Hypnosis (AJCH) are evaluated, and suggestions for upgrading the value of case study designs are discussed.

FRAMEWORK FOR EVALUATION

Internal and External Validity

Bloom & Fischer (1982, p. 290) state: "From the point of view of the canons of logic, case study methods are grossly lacking in power to make causal inferences . . . ." The validity of claims of causal connection between an intervention and observed change is referred to as *internal validity* (Bloom & Fischer, 1982). The internal validity of practice efforts, that is, the validity of claims that a specific intervention caused a change in a clinical problem, must be of great importance to clinicians. The development of potent and reliable therapeutic technology depends upon valid connections between treatment and outcome.

Kazdin (1982, p. 92) lists five primary threats to internal validity in the study of individual cases: history, maturation, testing, instrumentation, and statistical regression. These will not be discussed in detail here but the reader is encouraged to become familiar with these concepts.

*External validity* concerns the replicability and generalization of the case study to other clients, problems, and practitioners (Bloom & Fischer, 1982). Replicability of the intervention used will be the primary concern in the framework to be developed below. An intervention must be replicable to be generalizable to other clients, problems, and/or practitioners.

Single System Designs

Single case experimental designs are methodological approaches allowing for greater confidence in statements concerning change and causal inference than traditional anecdotal case studies (Hersen & Barlow, 1976). Bloom and Fischer (1982) note single case designs are:

> . . . fundamentally different from the case study in a number of important ways, including the fact that single system designs involve planned use of a research design, clear measurement rules, explicit evaluation procedures, and a clear identification of an intervention program, including when intervention starts and when it is completed. (p. 294)

Bloom and Fischer (1982, p. 240) further note two primary purposes in using single system designs. The first is monitoring problems continuously to determine
whether or not a problem actually changes. This is done through use of repeated, objective measures. The second endeavor involves attempting to determine any causal connection between an intervention and observed change. Both these endeavors concern internal validity, and single system designs offer greater power than traditional case studies in providing valid conclusions (Hersen & Barlow, 1976; Kazdin, 1982; Bloom & Fischer, 1982). Thus, single system designs should be used whenever possible.

Though single system designs represent a powerful methodology, there are a number of problems in their implementation (Thomas, 1978) and it is not always possible nor desirable to use them. If certain design characteristics are used, however, the case study provides a viable alternative to use of single case experimental designs. A framework, which highlights desirable design characteristics, was developed for evaluating case studies, and was then used to evaluate the articles reviewed in this study. Based on this study, suggestions are presented concerning use of case study methodology.

The Framework

Kazdin (1982, pp. 87-102) discusses criteria that can be used to assess the internal validity of case studies. These criteria can be seen as characteristics which, by their collective presence, can make the case study approach the single case experimental design in internal validity. These criteria are: type of data, assessment occasions, past and future projections of problem level, and number and heterogeneity of subjects. Each of these dimensions is discussed briefly.

Type of Data

Type of data refers to the data or information that is used to determine whether change has occurred. This dimension, called the dependent variable in research terms, refers in effect to the definition of the clinical "problem" and how it is measured. The type of data can range from anecdotal information to objective data, such as standardized self-report measures (the Beck Depression Inventory, for example), direct measures of overt behavior, or ratings by other persons. Kazdin (1982) states, concerning the internal validity of case studies:

...the absence of objective or quantifiable data usually precludes drawing conclusions about whether change occurred. (p. 91)

Thus, anecdotal evidence, without use of objective measures, results in inferences concerning change and causality reflecting merely the therapist's opinion (Kazdin, 1982, p. 88).

Objective data result from some measurement of the presenting clinical "problem." Being able to measure the "problem" presumes the "problem" has been well defined. Bloom and Fischer (1982, p. 34) note, "... if a problem exists, it can be measured." They go on to present a plethora of measurement procedures that can be used to measure any "problem" confronted by the clinician. Thus, the first step in improving the internal validity of case studies is the development of an operational definition of the "problem" to be treated. This definition should lead directly to an objective measurement procedure.

An example of a study using objective measures is provided by Gwynne, Tosi, and Howard (1978). This study, essentially an A-B-C single case design, uses three measures of "non-assertion:" a self-report measure of "anxiety," a self-report measure of "subjective units of discomfort," and a self-report of number of approaches.
to people. In contrast, an example of anecdotal data can be seen in a case study by Pelletier (1977). The problem, "hysterical aphonia," experienced by a 55-year-old woman, is defined and measured anecdotally (p. 149): "Frequently she could not speak at all. At other times, she labored to whisper only a few words. She panicked when the phone rang while she was alone at home." This description serves as the only problem definition and measurement.

In light of the above, the first dimension on the evaluation framework is use of objective data. This dimension is defined as: assessment of the "problem" to be treated in such a way that one or more objective measures of the "problem" are defined and used in the case study. This dimension, from the foregoing discussion, can be seen to be a necessary, though not sufficient, criterion for making claims of the occurrence of change and any causal connection between the intervention and change.

**Assessment Occasions**

Assessment occasions refers to the number and timing of objective measurements (Kazdin, 1982, p. 89). Major options are one or two time measurements (such as post-treatment only, or pre- and post-treatment measurements) or continuous, repeated measures over time. Kazdin (1982) notes:

When information is collected on one or two occasions, there are special difficulties in explaining the basis of the changes. Threats to internal validity (e.g., testing, instrumentation, statistical regression) are especially difficult to rule out. With continuous assessment over time, these threats are much less plausible, especially if continuous assessment begins before treatment and continues over the course of treatment. (p. 89)

Deiker and Counts (1980) provide a study exemplifying use of repeated measures before, during, and after treatment. Gwynne, Tosi, and Howard (1978) provide a second example. If a case study includes repeated measures of the "problem" over time, some of the major threats to internal validity can be ruled out. Use of repeated measures, thus, is an invaluable design feature and should be used whenever possible.

**Pre- and Post-Treatment Measurement Design**

Use of a pre- and post-treatment measurement design is the second dimension on the evaluation framework. This dimension is defined as: use of objective measure(s) of the "problem" to be treated once before and once after use of an intervention. Use of repeated measures is the third dimension on the evaluation framework. This dimension is defined as: use of objective measurement(s) of the "problem" to be treated repeatedly before and/or during, or after use of an intervention. Measurements taken prior to use of an intervention constitute the data baseline. Use of repeated measures both before and during use of an intervention constitutes a minimum requirement for an A-B single case design. A study by Welsh (1978) illustrates use of repeated measures during treatment. The studies by Deiker and Counts (1980) and Gwynne, Tosi, and Howard (1978) are examples of basic single case designs used with hypnotic interventions.

**Stability Information**

Past and future projections of problem level refer to, essentially, the stability, or trend, of the "problem." Kazdin (1982) writes:

Past and future projections refer to the course of a particular behavior or problem. For some behaviors or problems, an extended history may be evident indicating no change. If performance changes when treatment is applied, the likelihood that treatment caused the change is increased. (p. 89)

If a "problem" can be shown to have an extended history of stability, whether
stable at a fixed level or stable deterioration, future projections would likely predict continuation of the stable pattern. Stability information can be gathered by repeated measures over an extended period of time before intervention, by constructing a "retrospective baseline," or by gathering evidence of a historical pattern of stability. A "retrospective baseline" can be constructed by using a measurement method, such as a standardized self-report scale, and making a historical measurement (Bloom & Fischer, 1982; Hudson, 1982). The data are gathered retrospectively and constitute an objective history of the "problem."

Stability information is the fourth dimension in the evaluation framework. Stability information is defined as: evidence of an extended history, a retrospective baseline, or current repeated measurement pattern that shows a long-standing, stable, non-changing or worsening "problem." If an intervention is applied and the stable pattern changes, this suggests that the intervention rather than other factors is responsible.

Number and Heterogeneity of Subjects

Number and heterogeneity of subjects refers to the number of cases included in the study, and to the diversity of types of persons and problems treated with a particular intervention. Demonstrations with more than a single case provide a sounder basis for attributing change to the intervention used. Further, change demonstrated among several people who differ in clinical problems, age, sex, education, socioeconomic status, etc., enhances the probability change was due to the intervention. Kazdin (1982, p. 91) notes: "Essentially, with a heterogeneous set of clients, the likelihood that a particular threat to internal validity (e.g., history, maturation) could explain the results is reduced."

Number and Diversity

Number and diversity is the fifth dimension in the evaluation framework. This dimension is defined as: use of multiple cases, with heterogeneous clients with respect to problem type and/or demographic variables.

Classes of Study Design

These five dimensions provide a framework that can be used to assess the internal validity of case studies reported in the American Journal of Clinical Hypnosis. This framework will be used to differentiate published case studies into one of five classes. A class one study design is defined as one utilizing only anecdotal data. This is the weakest design. With no objective data, no basis for claims of change nor for making any causal inferences exists beyond the therapist's opinion.

A class two study design is defined as one using objective measures in a pre- and post-treatment measurement design. This provides two assessment occasions. If change is observed between the two measurements, then the use of objective data provides a basis for claiming change that is somewhat stronger than that provided by anecdotal data (Bloom & Fischer, 1982; Kazdin, 1982). However, threats to internal validity cannot be ruled out and causal inferences can only be speculated (Kazdin, 1982). Hence, this class design cannot be used if the practitioner wishes to draw sound, logically based causal inferences. However, this class design appears a minimum requirement for making claims of the occurrence of change that are based upon more than personal opinion.

A class three case study design is defined as one using objective measures of the "problem" repeatedly during treatment. Use of this class design provides a more sound basis for making claims of the occurrence of change. Studies using repeated
objective measures of the "problem" before and during treatment will be classed as single case designs. Use of repeated measures before and during treatment, essentially an A-B single case design, appears a minimum requirement for making data-based causal inferences (Kazdin, 1982; Bloom & Fischer, 1982).

A class four case study design is defined as one using repeated measures of the "problem" during treatment and having stability information. The presence of stability information constitutes past and future projections of the continuation of the "problem." Thus, the threats to internal validity related to assessment are further ruled out and history and maturation become less plausible. Thus, causal inferences from this class study are possible, though relatively weak (Kazdin, 1982).

A class five case study design is defined as one using objective measures repeatedly during treatment, stability information, and multiple cases. The addition of multiple cases helps further rule out history and maturation as rival hypotheses for change. Hence, this class study provides a stronger basis relative to the other four classes for making claims of change and causal inference (Kazdin, 1982).

The above five case study design classes can be considered pre-experimental designs. The class five design, even though it provides a relatively sound basis for claims of change and causal inference, and is superior to the other four, still lacks the power of single case experimental designs. The A-B design, with stability information, and the multiple baseline single case designs seem particularly well suited for use with hypnotic interventions. Their use, whenever possible, is strongly recommended.

REPLICABILITY OF INTERVENTION

One dimension will be used to assess the external validity of the published case studies: replicability of the intervention used. Three procedures of replication of single case results have been defined: direct, clinical, and systematic (Hersen & Barlow, 1976). These form, essentially, the heart of generalizing results of single case studies. Essential to any generalization endeavor is the ability to replicate the intervention used. Bloom and Fischer (1982) write:

... even though different independent variables are inevitable to some degree, practitioners seeking to extend results from one situation to another should attempt to make the interventions as similar as possible. (p. 262)

They also note (1982):

The independent variable in research is that which is presumed or expected to have an effect on the dependent variable. This has to be defined in specific measurable terms so that the researcher can know that it was indeed applied. So, too, in practice the independent variable must be clearly specified. In practice the independent variable is the intervention program — the strategy and specific techniques and procedures the practitioner will use to change the client or client/system. (p. 17)

The inability to replicate a specific intervention is listed by Bloom and Fischer (1982, p. 261) as a threat to external validity. Inability to replicate the intervention used in a successful case seriously limits the usefulness of the information for other practitioners.

Thus, the dimension to be used in assessing the external validity of the published case studies asks the fundamental question: Is the intervention used defined in such a way as to describe and specify all actions/behaviors of the therapist and/or client that constitute the intervention, and the specific order and timing of these actions/behaviors? This question asks, in effect, if enough operational information on the intervention is provided for an adequately trained practitioner reading the de-
TABLE 2

INTERNAL VALIDITY EVALUATION RESULTS

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1982/17</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>1981/22</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>1980/11</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>1979/14</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>1978/10</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td><strong>Total: 74</strong></td>
<td><strong>7</strong></td>
<td><strong>2</strong></td>
<td><strong>5</strong></td>
<td><strong>24</strong></td>
</tr>
<tr>
<td>% of Total:</td>
<td>9.46%</td>
<td>2.7%</td>
<td>6.7%</td>
<td>32.4%</td>
</tr>
<tr>
<td>No. Class 1 Studies: 67</td>
<td>No. Class 2 Studies: 2</td>
<td>No. Class 3 Studies: 2</td>
<td>No. Class 4 Studies: 0</td>
<td>No. Class 5 Studies: 0</td>
</tr>
<tr>
<td>% of Total:</td>
<td>90.5%</td>
<td>2.7%</td>
<td>2.7%</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

Number of Studies Using Single Case Designs: 3

**Results**

The above-defined framework for evaluating internal and external validity was used to assess 74 case studies published in the *American Journal of Clinical Hypnosis* from January, 1978, to January, 1983. The results are shown below in Tables 2 and 3.

**Internal Validity**

Of the 74 studies evaluated, 67 used no objective measures (90.5%). These 67 are class one designs. As noted earlier, a class one design furnishes no information concerning change or causality beyond the opinion of the study's author.

Two studies utilize objective data in a pre- and post-treatment measurement design. These are class two studies. Specifically, these are studies reported by Boutin (1978) and Horsley (1982). A class two study provides a data basis for claiming the occurrence of change. However, the pre- and post-treatment measurement design is not strong enough to rule out threats to internal validity and, hence, claims of change have a weak logical basis and causal inference remains a matter of surmises. These two studies represent 2.7% of the studies reviewed.

Two studies utilize objective data in a repeated measures design. These class three designs were published by Welsh (1978) and Godec (1980). A class three design
provides a more firm basis for claims of change. Also, the repeated measures design tends to rule out threats to internal validity related to assessment. However, history and maturation remain plausible alternative explanations for the data and, hence, causal inferences are somewhat weak. These two studies represent 2.7% of the total studies reviewed.

There were no class four or five studies found among the studies evaluated. Three single case designs were found among the studies reviewed. Gwynne, Tosi, and Howard (1978), and Deiker and Counts (1980) utilize A-B-C single case designs in their studies. Deiker and Counts (1980) refer to their design as a multiple baseline design. However, this is not accurate. Their design is best described as an A-B-C design using multiple dependent measures. A multiple baseline design employs repeated measures of either different behaviors or persons with the intervention applied in a temporally staggered pattern across behaviors or persons. The third single case design found among the studies reviewed was published by Lewis (1979). This study employed an A-B-A-B single case experimental design. These three studies represented 4.05% of the total reviewed.

**DISCUSSION**

The implications of this study are clear. In about 91% of the studies reviewed there is no information concerning change beyond the practitioner's opinion. If a class two study is taken as the minimum allowable for making data-based claims of the occurrence of change, then only 7 of 74 (9.5%) of the studies reviewed allow for claims of change. Further, if a class four or five design, or at least an A-B single case design, is taken as a minimum requirement for making causal inferences, then causal inferences can be made in only 3 out of the 74 case studies reviewed (4.05%). Even if a class three design (a very weak design for making causal inference) is taken as a minimum for causal inference, then such inferences can be made in only 5 of the 74 cases reviewed (6.75%). In only 10 of 74 case studies is the intervention replicably described (13.5%). If the use of hypnosis clinically is to be reliably developed, these deficiencies must be corrected.

It is suggested that practitioners use single case designs whenever possible. As mentioned earlier, the A-B design, used with stability information, and the multiple baseline design seem particularly well suited for use with hypnotic interventions. It is thus recommended that practitioners using hypnotic interventions become familiar with these designs.

If single case designs are not used, then it is suggested first of all that practitioners state explicitly the purpose of their case studies. If the practitioner wants to present a case in which, in his/her opinion only, change occurred that might have been due to a particular intervention, then a class one design is appropriate. A class one design can be used to alert fellow practitioners to the possibility that change occurred in a particular case and to the

**TABLE 3**

<table>
<thead>
<tr>
<th>Class A Designs</th>
<th>Class B Designs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number: 10</td>
<td>Number: 64</td>
</tr>
<tr>
<td>% of Total: 13.5%</td>
<td>% of Total: 86.5%</td>
</tr>
</tbody>
</table>

**External Validity**

As Table 3 shows, 64 of the 74 studies (86.5%) were class B studies. In these studies, the interventions used were not described in a clear, replicable manner. Only 10 of the 74 studies (13.5%) were class A.
possibility that a given intervention played a causal role. The limitations of the class one design should be kept in mind by both case study publishers and readers. If claims of change are to be made, a class two design should be a minimum requirement. If causal inferences are to be made, a minimum requirement would seem to be either class four or five designs or at least an A-B single case design. Whenever possible, objective data in a repeated measures design, stability information, and multiple, heterogeneous cases should be used. These design features can greatly enhance the internal validity of case studies. It is suggested by the author that a basic methodological design feature of case studies be use of repeated measures of the "problem" both before and during treatment. With this design feature the clinician can make sound, data-based claims of change and can begin to consider causal inferences. This will ensure a basic level of internal validity in the study. This is crucial to the development of reliable hypnotic technology.

Final suggestions concern operationally defining the intervention used. This should be done in such a manner that any adequately trained professional reading the study can replicate the intervention. This is essential to replication efforts and the utility of published case studies for other practitioners. It is in this direction that the greatest challenge for the hypnotherapist lies. Operationally defining such constructs as "hypnotic trance depth," "hypnotic suggestions," and other components of hypnotic interventions may be a prerequisite to developing a reliable and valid hypnotic technology. Further, new conceptual procedures may need to be developed to successfully put into operation the person-centered techniques of Ericksonian "utilization approaches." It is suggested any definitions of interventions be made from the perspective of a fellow clinician attempting to replicate the intervention used.

Conclusions
Authors such as Kazdin (1982), Bloom and Fischer (1982), and Hersen & Barlow (1976) believe information developed from studying the single case may be more valuable in developing effective therapeutic technology than traditional group comparison research. Doing this, however, will require sound single case methodology. Further, these authors suggest that clinicians are at the forefront in developing therapeutic technology. This article is intended to give an evaluation of current case study methodology as used by practitioners using hypnotic techniques and to give suggestions for upgrading this methodology. Concerning the current methodological state of case studies published in the American Journal of Clinical Hypnosis, the results of this review speak for themselves. Changes must be made if hypnotherapists are to be at the forefront in developing reliable therapeutic technology.

REFERENCES