DOCUMENT RESUME

ED 127 539 CG 010 767

AUTHOR Pipstra, Constance C.; And Others
TITLE Empirical Case Study: An Introductory Workshop.
PUB DATE Apr 76
NOTE 16p.; Paper presented at the annual meeting of the American Educational Research Association (San Francisco, California, April 19-23, 1976)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Behavior Change; Behavior Patterns; Case Studies (Education); *Change Strategies; Counseling; Data Analysis; *Educational Research; *Interaction Process Analysis; Intervention; Models; Observation; *Psychological Design Needs; *Research Design; *Statistical Analysis; Workshops
IDENTIFIERS *Median Trend Analysis

ABSTRACT

This speech presents the use of experimental designs and research procedures to help counselors develop effective treatment methods for clients. Baseline and treatment phases are delineated as the two basic phases occuring in most "N=1" designs. These designs are divided into the categories of single and multiple time series designs. The research procedures for the "N=1" designs are systematically described, followed by sections on data analysis and interpretation. The report includes references, a bibliography, and an appendix detailing the steps for a median trend analysis. This type of statistical analysis is recommended highly for nonstatisticians, because it requires a minimal amount of calculation. (HLM)

19.08

EMPIRICAL CASE STUDY:

AN INTRODUCTORY WORKSHOP

CONSTANCE C. RIPSTRA
Texas A & I University at Corpus Christi

PAMELA S. HIGHLEN
University of Western Ontario

NANCY L. VOIGHT
University of North Carolina, Chapel Hill

AMERICAN EDUCATIONAL RESEARCH ASSOCIATION

ANNUAL MEETING

SAN FRANCISCO, CALIFORNIA

APRIL 1976

I. Basic Phases of "N = 1" Designs

The two basic phases occurring in most "N = 1" designs are the baseline and treatment phases. The baseline phase consists of a series of observations without the application of any researcher-imposed conditions. The data collected in this phase provide an estimation of the pattern and frequency of the target behavior as it occurs naturally. The recommended procedure for collection of all data, and especially of baseline data, is through the use of unobtrusive measures or observations. This procedure prevents the baseline data collection process from actually being a treatment, since the mere process of observation may systematically change behavior and confound results.

The second phase is typically the treatment period consisting of a series of observations when a systematic intervention is in effect. With the exception of the passage of time, the addition of the treatment must be the only condition to vary from those conditions existing in the baseline phase. When this is the case, the differences between the baseline and treatment phases may be attributed to the imposed treatment.

The following symbols are widely used to designate various phases of intensive designs (Thoresen, 1972):

A = Baseline Phase

B = First Treatment Phase

C = Second Treatment Phase

D = Third Treatment Phase

F = Follow-up Phase

Regardless of the type of N = 1 design chosen, the counselor-researcher will often include a follow-up phase. This phase provides a check on the target behavior after the experiment has been completed to see if behavior change has been maintained.

II. Types of Intensive Designs

The two general types of intensive designs are the single and multiple time series designs. In the single time series design only one baseline is taken on one target behavior, although more than one treatment may be employed. In contrast, the multiple time series design involves two or more baselines taken concurrently across situations, behaviors, or individuals.

Single Time Series Designs. The simplest of these designs is the A Design. An A design is used to examine natural changes over time when no systematic intervention is employed. Because no treatment is involved, this design is primarily descriptive. The counselor-researcher typically notes long term trends in data and any fluctuations about these trends.

The AB design consists of a baseline and treatment phase. After data collection, the counselor-researcher compares the frequency of the behavior during both phases. If desired change in the behavior occurs, one can only conclude that the intervention seemed to be effective.

The third category of single time series designs involves replication of the baseline phase. Replication of the original treatment as well as the introduction of new interventions is also possible. Through replication of baseline and treatment phases, the investigator may establish experimental reliability if the results of the first two phases are replicated. Hence, replication single time series designs are preferable over the AB designs. Examples of such designs involving replication include: ABA, ABAB, ABACA, and ABCDEAE.

However, single time series designs have some limitations. First, the counselor-researcher is limited to examining one treatment or a sequence of treatments and one behavior over time. An additional problem is that the behavior being studied may not be reversible Many intervention strategies involve teaching the client new skills, such as acting more assertively. Many social and intellectual skills cannot be forgotten once the treatment is removed. Hence, the change in behavior may be perpetuated after the intervention is withdrawn making the establishment of experimental reliability impossible. Also, the counselor may find it unethical or undersirable to withdraw the intervention. In many cases treatment programs should be designed to insure that desired change is maintained. Finally, if time for the study is limited, the counselor may not be able to introduce replication phases or additional treatments. These weaknesses may be overcome if the counselor uses a multiple time series design.

Multiple Time Series Designs. The multiple time series design requires that concurrent baselines be taken across situations, behaviors, or individuals, so that several baselines are established before an intervention is introduced. The treatment is then applied to only one situation, behavior, or individual. If a change is observed at the onset of treatment, then a causal inference may be made between the change and the treatment effect. The first type of multiple time series design involves baselines taken across situations. Counselors are often times interested in whether clients change that occurs in one setting (e.g., the classroom) also takes place in another situation (e.g., the home). For the client who is attempting to change more than one behavior a multiple time series design across behaviors would be chosen. If the counselor is working with several individuals with similar concerns, multiple baselines across individuals could be utilized.

As with the single series design, the multiple time series design provides a within-client comparison since the individual serves as his own control. In addition, ongoing events in the client's life are controlled for because multiple baselines are taken. A pre- and post-treatment comparison for each behavior, individual or situation under study is possible. The counselor-researcher is not limited to studying one behavior in one situation. Hence, these designs afford the counselor greater flexibility. Finally, the multiple time series designs avoid ethical and treatment problems of removing the intervention in order to assess its effectiveness.

The counselor's choice of an appropriate intensive design depends on the specific nature of the client's concern. However, the counselor-researcher is encouraged to carefully consider the merits of the multiple time series designs since they overcome the inherent limitations of single time series designs.

III. N = 1 Research Procedures

The procedures to be followed for performing an "N = 1" study are consistent regardless of the design used and are summarized as the following steps:

1. Precisely define the behavior(s) to be observed. The behavior can be described in gross or fine terms as needed for meaningful results. However, the observer must be able to distinguish between the presence or absence of the behavior. An optional prior step which may be helpful in defining a behavior is to observe the client in the setting. This observation can help distinguish the most important behavior(s) to be acquired or eliminated. For modification of complex behavior it is advisable to take one key aspect of the behavior.

2. Decide on the details of the data collection. The length of the observation period per day, the structuring of that observation period, and when the observations are to be made are the major considerations. An observation period typically is divided into small units which are alternately observation and recording times. For instance, for a 20-minute daily observation each unit may be 15 seconds. The first unit would be observation, the second unit would be recording what was observed in the first unit, the third unit would be observation, and so on. Using this particular system, there would be a total of 40 observation units per day. However, the details of data collection depend on the behavior to be observed. The objective is to establish a procedure to collect accurate and meaningful data. The decision of when the observations are to be made should take into consideration the client's schedule. This observation time should be the same throughout the study.

It is important to note that it be impossible for the frequency of the behavior to have an attainable ceiling or floor. In other words, the units of observation should be small enough or the definition of behavior specific enough to avoid a 100% or 0% frequency count for any one observation time.

3. Specify the number of observations per phase or length of the phases. The phases in the design need not have an equal number of observations, but a

minimum of seven to nine data observations per phase is recommended to permit adequate prediction beyond the data collection period (White, 1972b).

4. Train the observer or data collector according to the behavioral definition in step #1. Training should include observer reliability checks, in which at least two people independently observe the same subject and the same behavior. The comparison of the independent results can determine that accurate data is being collected.

5. Keep a diary to record the uncontrollable conditions and events which may influence the target behavior. Have the client keep one also. Such infomation can be used to hypothesize about possible causes of atypical frequencies of the behavior.

6. As data collection proceeds, plot the data points on a graph, with frequency or rate of the target behavior on the vertical axis and observation days or time on the horizontal axis. Draw phase change lines, lines perpendicular to the horizontal axis, to distinguish between the phases of the study.

IV. Data Analysis

The process of statistical analysis is less essential for the empirical case study than it is for group research. A definition of a meaningfully significant difference between pretreatment and post-treatment behavior levels is the relevant evaluation of the treatment. Therefore, the questions to be asked are whether the behavior change is in the desired direction and whether the magnitude of change is sufficient to satisfy the objective of the counselor and client. The demonstration that these have been answered satisfactorily is the relevant evaluation of the treatment's effectiveness.

Procedures do exist for performing more rigorous analysis than the "eyeballing" procedure just mentioned. Dr. Owen White (1972 a,b) has developed an analysis procedure based on the median of the data as the measure of central tendency. Median statistics is a new approach to data analysis and is being recommended as the best available method for nonstatisticians, because it requires a minimal amount of calculation. White (1972b) has discussed the merits of the use of a trend line based on the median. The main argument in favor of median statistics versus those based on the mean is that deviant data points are disproportionately weighted in the computation of the mean. All data points have equal weighting in the computation of the median. Therefore, a median score is more representative of data when there is one or more deviant scores in the data. Detailed instructions for analyzing data may be found in White's Working Paper No. 16 (1972b), which is summarized in Appendix A.

V. Data Interpretation

Specific interpretation of the data is dependent on the design and the type of data analysis utilized in the study. However, general guidelines of interpretation for all "N = 1" studies do exist.

One guideline is that the results are not generalizable beyond the client/subject unless the results have been replicated across a number of subjects. Therefore, ordinarily the researcher may not legitimately draw conclusions about the general effectiveness of the treatment. However, the reader of an "N = 1" study may make such generalizations if the researcher has sufficiently described the client/subject. This is allowable according to the "Cornfield-Tukey Argument" (Cornfield & Tukey, 1956) which concludes that results from a nonrandomly selected sample may be generalized beyond that sample by the reader when a sufficiently complete description of the sample is reported.

This type of study "... also precludes the use of 'explanatory terms' based on hypothetical mental and psychological states and variables ..." (Bijou, et al, 1969, p. 209). Interpretation should consist of a description of the results in relation to the specific procedures and treatment and appropriate comparisons with the results of other studies with similar conditions and procedures.

The following chart will be of assistance in the interpretation of the results of an "N = 1" study.

KEY: G = graphing

MB = median trend line and benomial test

a - were objectives met?

b - what changes occured?

c - was treatment responsible for the changes?

d - could this have happened by chance?

e - if more than one treatment were used, which was more effective?

f - can this treatment be used with other people, situations, behaviors?

Design	Data Handling	Results Possible
A	G	a,b
AB	G	a,b
AB	MB	a,b,d - some hints that c may be true
reversals	G	a,b - some hints about c and e
reversals	MB	a,b,c,d, (e)
multiple baseline	G	a,b,f - hints about c
multiple baseline	MB	a,b,c,d,f
mixed	G	a,b,f - hints about c
mixed	MB	a,b,c,d, (e) f

References

Bijou, S.W., Peterson, R.F., Harris, F.R., Allen, K.E., & Johnston, M.S. Methodology for experimental studies of young children in natural setting. The Psychological Record, 1969, 19, 177-210.

Cornfield, J., & Tukey, J.W. Average values of mean squares in factorials. Annals of Mathematical Statistics, 1956, 27, 907-949.

Thoresen, C.E. The intensive design: An intimate approach to counseling research. Paper presented at the Annual Meeting of American Educational Research Association, Chicago, 1972.

White, O. Working paper no. 15: The prediction of human performance in the single case: An examination of four techniques. Eugene, Oregon: Resource Center for Handicapped Children, 1972a.

White, O. Working paper no. 16: A manual for the calculation and use of the median slope -- A technique of progress estimation and prediction in the single case. Eugene, Oregon: Resource Center for Handicapped Children, 1972b.

Bibliography

Bijou, S.W., Peterson, R.F., Harris, F.R., Allen., K.E., & Johnston, M.S. methodology for experimental studies of young children in natural settings. In Thomas, E.J. (ed.) Behavior modification procedure.: A sourcebook. Chicago: Aldine, 1974.

Huck, S.W., Cormier, W.H., & Bounds, W.G. Reading statistics and research. NY: Harper & Row, 1974, Chapter 15.

Browning, R.M., & Stover, D.O. Behavior modification in child treatment. Chicago: Aldine, 1971, Chapter 4.

White, O.R., & Liberty, K.A. Behavioral assessment and precise educational measurement. In Haring, N.G. & Schiefelbusch, R.L. Teaching special children NY: McGraw-Hill, 1976.

APPENDIX A

MEDIAN TREND ANALYSIS

The steps for a median trend analysis are as follows:

1. Determine the median data point by rank ordering the data points. If there is an odd number of data points, the median will be the middle point. If there is an even number, use the middle number that is closest to the center of the distribution on the x-axis.

   For example: for data 1 2 5 9 11 the median is equal to 5

   for data 1 2 5 9 11 12 the median data point is either 5 or 9 depending on which is closest to the middle of the graphed data.

2. Determine the median trend line by following these steps:

   a. Number the data points on the graph according the the x-axis (see Figure 1).

   b. Using the median data point (step #1) as an axis, rotate a ruler from the vertical position in a counter-clockwise direction. Note the order of the data points which are crossed by the ruler on <u>both</u> sides of the axis. In this way all data points should be ordered by a 180^{o} arc of the ruler.

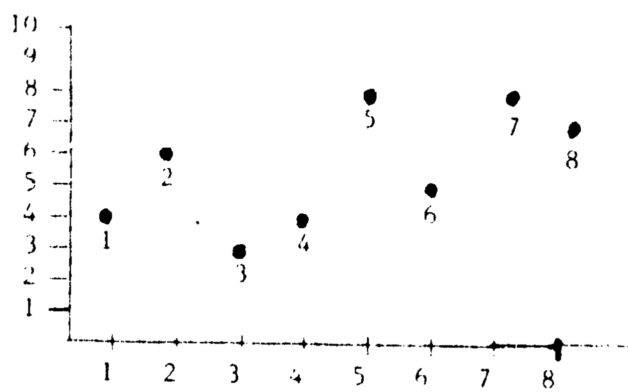

Figure 1

   (1) Since there is an even number of data points, the median is either #2 or #6. #6 is closer to the center of the data distribution and therefore is the median point to use.

   (2) Following step 2b, the order would be (6) as the axis
5
2
1
4
8
3
7

c. Subtract the position point of the median (axis) point from each of the ordered points. Use absolute values.

(6)	(step c)	
5	5-6=	1
2	2-6=	4
1	1-6=	5
4	4-6=	2
3	3-6=	3
8	8-6=	2
7	7-6=	1

d. Cumulatively add those values derived from 2c above.

(6)			(step d)	
5	5-6=	1	1	
2	2-6=	4	5	
1	1-6=	5	10	
4	4-6=	2	12	
3	3-6=	3	15	
8	8-6=	2	17	
7	7-6=	1	18	(cumulative total)

e. Divide the cumulative total by 2.

18/2 = 9

f. Scan the cumulative summation column until that point is reached or just exceeded. Note the point with which that cumulative total point is associated.

(6)		(step d)	
5		1	
2		5	
1	←---	10	** (10 > 9)
4		12	
3		15	
8		17	
7		18	

g. The above results indicate that a line connecting points #6 and #1 will be balanced. To check that this indeed is the median trend line, steps b. through f. must be repeated using data point #1 as the axis. This process is repeated until a pair of points check out for each other.

(1)		c		d	
3		2		2	
4		3		5	
6		5		10	
8	←--	7	←-	17	** (17 >14)
7		6		23	
5		4		27	
2		1		28/2 = 14	

(8)		c		d	
7		1		1	
5		3		4	
2		6		10	
1	←--	7	←-	17	** (17 >14)
4		4		21	
3		5		26	
6		2		28/2 = 14	

For this example the trend line of #6 to #1 did not check out as the best trend line. However, upon continuation of the process, a trend line of #1 to #8 did check out as was illustrated above.

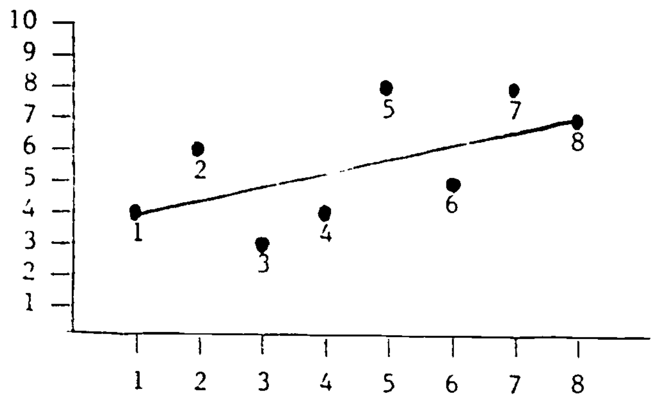

Figure 2

If the data are particularly stable or closely bunched, this method may prove inadequate since it will be difficult to determine which of two or more data points are intersected first. In such cases large charts can be constructed to spread out the data.

When small amounts of data are used, there is a reasonable probalility that the data will be so evenly balanced at any given point in time that there will actually be two different possible median trend lines. When the number in the cumulative sum column (step 2c) is equal to (rather than greater than) ½ of the total sum, it is a signal that the data are equally balanced, and that there are two possible median trend lines which must be tested. This situation is called branching.

Consider the following data:

```
(5)
4    4-5=  1    1
3    3-5=  2    3
1<--1-5=   4<---7 ** (7> 5)
2    2-5=  3    10/2= 5
```

```
(1)
2    2-1=  1    1
5<--5-1=   4<---5 ** (5=5)
4    4-1=  3    8
3    3-1=  2    10/2= 5
```

In this case one median trend line has been confirmed (i.e. a line running between data points 1 and 5). The second point to be tested in a branching situation will always be the next one down in the column listing the order in which the data are intersected. In this case data point #4 will be used as the next axis.

```
(4)
5    5-4=  1    1
3    3-4=  1    2
1<--1-4=   3<---5 ** (5>3.5)
2    2-4=  2    7/2= 3.5
```

The data point associated with ½ of the cumulative sum total is #1, which was the last center of rotation. Therefore, the line between data points #1 and #4 is also a median line.

Which line, then, should be used? Whenever possible, collect more data. Usually only one more data point will be needed in order to establish a single median trend line. With the data given, however, there is no decision regarding the choice of one median trend line that can be made with complete confidence. White (1972b) suggests the following set of rules for deciding which line to use in further analysis:

1. If there is no particular hypothesis as to how the data should be progressing (or how one wishes it to progress), then select the line which is closest to being flat (i.e. showing no progress), since that will be the most conservative estimate in most cases.

2. If hypotheses are involved, or tests of significance are going to be run, then select the slope which minimizes the possibility of finding differences or supporting favorable hypotheses. By doing so, any statements of change will tend to be conservative.

The probability of more than one appropriate median trend line is relatively high with small amounts of data (4 or 5), but decreases sharply with greater numbers of data.

3. Determine meaningful and/or statistical significance.

   a. Meaningful significance should be determined as a sufficient improvement in behavior as defined by the learning objective and should be considered separately from statistical significance. The general aim would be that some intervention changed the predicted frequency or behavior from baseline in the desired direction.

   The steps for evaluating the data for meaningful significance are as follows:

   1) Using a dotted line, extend the trend line of one phase into the next phase. The dotted line is the predicted rate of behavior.

   2) Visually determine the step and the progress change between the predicted rate of behavior and the actual trend line for the second phase. The step is the difference between the actual and predicted trend lines at the beginning of the second phase (see Figure 3). The progress change is the difference between the slopes of the predicted and actual trend lines. Both the step and progress change estimates define the total difference between predicted and actual behavior. Even if there is a significant step between phases, if the slopes are such that in time the behavior will be at the same level regardless of the intervention, the overall effect will not be meaningful (see Figure 4).

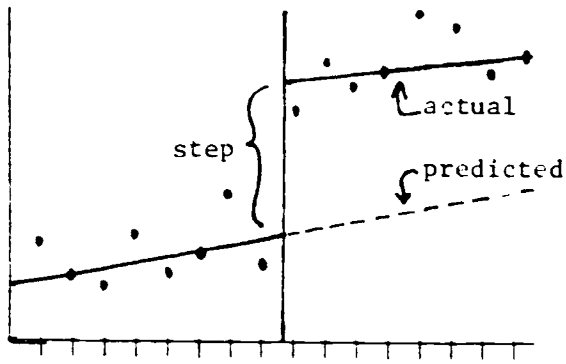

Figure 3

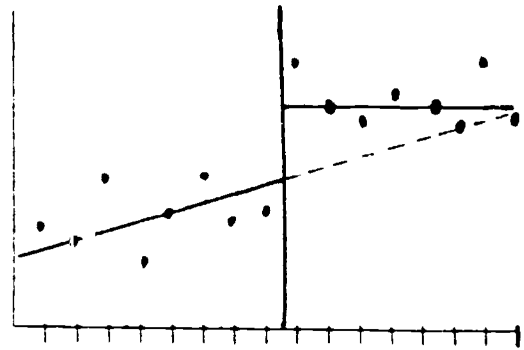

Figure 4

b. Statistical significance is calculated as follows:

1) As with meaningful significance the first step is to extend the trend line of one phase into the next phase with a dotted line.

2) Note the number of data points which fall on either side. The null hypothesis is that if there are no differences between the phases, an equal number of data points should fall on each side of the predicted trend line. In statistical terms H_o: $p = .5$.

3) The statistical test is a binomial test:

$$P(r) = {}_{N}C_{r}\ (p)^{r}\ (q)^{N-r}$$

Where P is the probability of ( )
r is the number of data points on one side of the predicted trend line
p is the hypothesized probability (.5)
q is 1 - p (.5)
N is the total number of data points in the second phase
C indicates a combination is computed for N and r

Example: In Figure 4 one data point falls below the predicted trend line. Therefore $r = 1$
$N = 8$
$p = .5$
$q = .5$

$$P(r = 1) = \begin{matrix} 8 \\ 1 \end{matrix}\ (.5)^{1}(.5)^{8-5}$$

$$= 8\ (.5)\ (.0078)$$

$$= 8\ (.0039)$$

$$= .0312$$

4) The decision must be made whether P(r) is small enough to say that the actual trend line from the data in the second phase is not a chance occurrence. Traditionally an alpha level of .05 is selected. In that case the probalility of having only one data point on one side of the predicted trend line (p = .0312)

is small enough for us to conclude that there is an overall difference between the two phases.

The above sequence is continued to determine differences between all consecutive phases. It is also used to analyze for differences between nonadjoining phases, such as might be needed for analyzing complex designs.