

excerpts from

Rival Hypotheses: Alternative Interpretations of Data Based Conclusions by Schulyer W. Huck & Howard M. Sandler

Copyright © 1979

Part TWO

1. Camping Out

Based upon the books you've read, the movies you've seen, the relatives you may have visited, or your own first-hand experience, you are undoubtedly aware of the fact that the daily routine in psychiatric hospitals can be stifling for the patients. Yesterday, today, and tomorrow can be indistinguishable, with weeks, months, and years blending together. Fortunately, there are some staff members who are sensitive to the ill effects brought about by the same routine day in and day out—and who care enough to try to do something about it.

Usually, these attempts at breaking the monotony are not evaluated by means of any sort of formal research investigation; consequently, no one really knows for sure whether the new activity accomplishes its desired goal. But on occasion, data are collected in an attempt to verify scientifically the worth of the innovative program. One such research study was conducted in Utah, and the new activity—camping out—was about as different from the daily institutional routine as you could imagine.

The subjects in this investigation were 25 men and women from age 19 to 62, who were randomly selected from a state psychiatric hospital located in an urban area of Utah. These individuals were taken to an isolated camp site in the mountains near Flaming Gorge. The patients and staff camped out for five days (Monday through Friday), and, while on their camping trip, the staff maintained a very low profile. The patients had the responsibility of forming teams for cooking and clean-up, of arranging sleeping accommodations, and of structuring their own free-time activities. Other than busing the group to and from the campsite, the staff took charge on only two occasions—when the group went on a raft ride down the river and when they visited a nearby store for snacks.

The researchers expected this week-long camping retreat to serve as a therapeutic tool, and in particular they hypothesized that the activities would lead to increased social interaction among the patients. To test this hypothesis, two types of data were collected on both the first and final days of the camp-out. According to a pre-arranged random-time sampling scheme, five-minute sessions of group interaction were taped, unobtrusively, on an audio recorder. In addition, photographs were taken of the patients.

One week after returning, all staff members and five of the patients who had gone on the camp-out used the audiotapes and pictures to rate the 25 patient campers in terms of social interaction. These ratings were obtained by using a modified version of the Bales Interaction Matrix. For each of the 25 patients, average ratings from the staff judges and the patient judges were computed. Then, the ratings within each group of judges were averaged across the 25 patients to obtain overall Monday and Friday ratings for the entire group. Since the Bales Interaction Matrix yields 12 subscale scores ("Gives suggestions," "Releases tension," and so on), there were two sets of 12 pre-test and post-test composite ratings on the 25 campers, one set from the staff judges and the other set from the five patient judges.

When the pre-test and post-test data were tested statistically, the researchers found that there was significantly more social interaction at the end of the five-day camping excursion than there had been at the beginning. The ratings from the patient group of judges showed increases on 11 of the 12

subscales of the Bales instrument, while the ratings from the staff members indicated significant improvement on all 12 subscales. One possible interpretation of these results is that the camping activities and unique environment brought about increased social interaction. **Might there be other plausible explanations for the observed differences between the Monday and Friday ratings?**

.....

2. Psychotherapy Revisited

One of the main plausible rival hypotheses to any study of psychotherapy is that of statistical regression. Individuals tend to enter therapy at an extreme point in their lives and, thus, tend not to be at such an extreme later on. This would generally be the case even in the absence of therapy. We are not criticizing the effectiveness of psychotherapy, only pointing out the difficulty of doing research in this area. Under these conditions, the need for studies that follow up psychotherapeutic programs should be obvious.

One such study investigated the effects of four psychotherapy programs, including one control condition, on the following three diagnostic groups: non-psychotic patients, short-term psychotic patients, and long-term psychotic patients. The researchers compared the results of a six-month follow-up with the results of the original study, as well as those from an 18-month follow-up of the same patients. The data consisted of responses to a nine-item questionnaire. Among other topics, the questionnaire addressed employment, re-hospitalization, and general adjustment. Data were missing for 10 of the 96 participants in the original study.

At the end of six months, significant group differences were found in the areas of degree of illness, employment, having friends, and remaining out of the hospital. These differences continued to exist at 18 months, although other earlier differences in community adjustment had disappeared by this time. However, significant differences in employment between the three groups that had existed at six-months ceased to exist by the time of the latest follow-up.

The researchers concluded that "the disappearance of significant employment differences by 18 months suggests that psychotherapy effects are of short-term duration" (Fairweather & Simon, p. 186). They also recommend the development of a social support system in the community for people similar to those in their study. **While we might agree with their recommendation, we are not sure the data support it. Are you? What alternative hypothesis might also explain these differences?**

.....

3. Newspaper Advertising

Each year, millions of dollars are spent on newspaper advertising. [Students: Recall that this book was published in 1979, when newspaper advertising was like website advertising.] Obviously, the people who pay for the ads feel their financial outlay is worthwhile. In other words, they believe that there is a cause-and-effect relationship with newspaper advertising being the cause and subsequent increased sales being the effect. But do the ads really bring about, in a causal sense, more consumer purchases? Two researchers conducted a study designed to answer this simple yet important question.

The setting for this study was a small town in northern Illinois (population about 3000). The subjects were 142 women who regularly purchased items at a grocery store. A list 28 items appeared in the local newspaper for four consecutive days prior to the day of the study. Each of these 28 products was advertised at a reduced price, and the purpose of the study was to determine whether this advertising made a difference in sales of these items.

Following the four days of advertising, the data of the experiment were collected. The procedural aspects of the study on this fifth day were as follows. As each participant came through the checkout counter, the clerk examined the participant's purchases to see if any of the advertised sale items

were included. If one or more of the 28 items were about to be bought, the clerk was instructed to ask whether the consumer had read about the sale items in the newspaper ad.

Although several of the participants came back to the store a second (or third) time on the same day that the data were collected for the study, responses from each participant were recorded only for her first time through the checkout counter. The results indicated that all 142 participants purchased one or more of the 28 advertised items. Ninety-nine of the participants stated that they had read about these items in the newspaper, while 43 participants admitted that they had not.

Based on these figures, the researchers concluded that "reading the newspaper advertising seemed to increase purchase of advertised items more than not reading the paper" (Peretti & Lucas, p. 693).

Based on the data that were collected in this study, can we conclude that a cause-and-effect relationship has been established? Does newspaper advertising lead to increased purchasing behavior? Or is there an alternative hypothesis we need to consider?

.....

4. Psychotherapy Revisited Again

Earlier in this volume, we discussed both the importance of and some of the problems with a study that examined the long-term effects of therapeutic intervention. Given the popularity of psychotherapy among college students, we didn't think that you would mind a second study in the same area.

Using a variety of psychological instruments, two psychologists in Chicago tested 93 college students before they entered therapy and then at the end of therapy. Eighteen months after the second testing, they were able to locate and test 69 of the original group. Comparisons made between the post-therapy and follow-up scores showed no significant differences, thus leading the researchers to conclude that the students did not continue to improve in the post-therapy period. This result was in contrast to the expectations of the psychologists, who believed that psychotherapy sets the stage for future growth on the part of the client.

Although any improvement from the pre-therapy to post-therapy testing might be a consequence of regression toward the mean, this would not be a factor in a change (or lack of change) in the post-therapy period. **What other rival hypotheses are there to the finding of no difference?**

.....

5. Alcoholics in Control

Most of us know people who recognize that they drink too much. The typical comment of these people is often "I could quit tomorrow if I wanted to," but somehow they do not ever want to. Many studies involving men alcoholics report that they really do believe that they are in control of their behavior, including their drinking. This led two researchers to study the question of control in women alcoholics.

Two scales were used to measure general control orientation and drinking control orientation in a total of 90 women. Data for the alcoholic group were collected on 50 women in two halfway houses over a six-month period (necessitated by the small population and low turnover of the halfway houses). A control group was used, which consisted of 40 volunteer social drinkers who were businesswomen; many were chapter officers of a national business women's association. The women in the control group did not differ significantly in either age or education from the women in the alcoholic group.

Statistical tests showed that the women in the alcoholic group were more external on both scales. This is in contrast to men alcoholics, who show greater internality and, thus, greater perception of control over their behavior. The authors gave two possible interpretations of their findings: Men and women alcoholics differ in their control orientation, and women alcoholics in halfway houses may not be representative of all women alcoholics. **Are there other possible explanations for the findings?**

6. To B Mod or Not to B Mod

One of our neighbors recently instituted a system using gold stars as reinforcers for her 7-year-old son in order to increase the time he spent reading and decrease the time he spent engaged in unacceptable behaviors. Because the neighbor's, older, 9-year-old son also wanted to be involved in this family project, there was little choice but to include him. The outcome was what any cynic would have predicted: The older child read more and behaved even better than usual, while the younger child gave up completely in the face of the competition. As their mother bemoaned the abject failure of behavior modification to modify the behavior of the child whose behavior need modifying, we were tempted to give her a copy of the study described below.

This study compared a behavior modification treatment technique and a placebo treatment. The participants in the study were twelve families, each of whom had more than one child. In each family, the target child was a boy (between 5 and 14) who had been engaging in one or more highly undesirable behaviors (such as setting fires or stealing). Half the families were randomly assigned to the behavior modification group, and the other half were randomly assigned to the placebo group.

The treatment group received training from a programmed text, in addition to a weekly group session in the presence of two experienced therapists. The placebo group did not receive training from a programmed text so they spent this time making tape recordings of their problems. However, they did have weekly group sessions with two therapists, just like the treatment group. Both the treatment and the placebo groups also received daily phone calls from the therapists.

A coding system was used by trained observers to describe each family's interactions (particular attention was paid to the deviant behavior of each family's son). All families' interactions were assessed before treatment started (pre-treatment assessment) and after treatment ended (post-treatment assessment). The observers who coded the families' behavior were unaware of the families' assignment to either the treatment or placebo groups. In addition, the parents completed a symptom checklist each week.

The amount of approval each parent received from the therapists during the group sessions was calculated, as was the amount of telephone-contact time each family received from the staff. Although both groups were supposed to receive equal professional contact, a significant difference was found in favor of the treatment group in the amount of telephone time received (118 minutes versus 57 minutes). As for the targeted deviant behaviors in the child, the treatment group showed a 61 percent decline (significant) while the placebo group showed a 37 percent increase (nonsignificant).

The authors concluded that the treatment was dramatically successful in reducing targeted deviant behavior. **But do you see any alternative explanation for the difference beyond the fact that the treatment group "received training from a programmed text"?**

7. Groups for Parents

Since we have young children who are occasionally less than perfect in their behavior, we are naturally attracted to studies that deal with behavior problems in children. "Groups for Parents" is a packaged method that offers parents both a support group of other parents and didactic information on a behavior modification program. The authors of "Groups for Parents" (along with a few others) published a study evaluating the effectiveness of their approach in "improving both general child behaviors [and] individually targeted ones." They also reported success in increasing the parents' rates of positive reinforcement along with the rates of compliance in their children.

The method of evaluation was quite simple. Thirteen groups of 20 - 25 parents (a total of 277 parents) met once a week for 2.5 hours over an eight-week period. About one-half of the parents had been referred by various community agencies; the rest had heard about the program from friends or other informed sources. The pre- and post-test measures included a problem behavior checklist

(that the parents filled out about their children), positive reinforcement rates (that the parents measured and recorded about their own rates of positively reinforcing their children's good behavior), compliance rates (that the parents recorded about how well they complied with the program), and client satisfaction (that the parents completed). Approximately two-thirds (180) of those enrolled completed the entire eight-week course.

The data analyses were equally straightforward, consisting of analyzing the differences between the parents' pre- and post-test data. Significant results were reported on the problem behavior checklist (parents reported that their children's problem behaviors decreased from pre-test to post-test), reinforcement rates (parents reported that they increased the number of times they gave their children positive reinforcement for the children's good behavior), and compliance rates (parents reported that they complied more closely with the behavior modification program after the treatment program than before). In addition, a very high rate of client satisfaction at the end of the study was reported. **On the basis of these results, should we enroll in these courses the next time they are available? Or do you see an alternative explanation for these positive outcomes?**

.....

8. Typed Papers and Grades

Most students like to receive high grades. Most parents like their children to receive high grades. Therefore, most students and parents would probably be willing to invest a modest amount of money in a service or commercial product if it were causally related to better grades. Some might even be willing to invest a great deal of money, depending upon how low one's grades have been in the past and how much value is attached to the higher grades that might be earned. Not surprisingly, college students have collectively paid thousands of dollars—possibly millions—to unethical businesses that write research or term papers for the students to turn in as their own work.

Grades are a powerful motivator. Realizing this, the people who market Smith-Corona typewriters [Remember: This book was published in the 1970s!] put out an advertisement in news magazines that implied that a student's grades are likely to go up if the student turns in term papers that are typed rather than handwritten. The ad was entitled "Students Who Type Usually Receive Better Grades," and the evidence came from a national survey of 400 high school and college instructors.

Each instructor was presented several statements and told to choose one of five responses for each statement: agree strongly, agree somewhat, have no opinion, disagree somewhat, disagree strongly. Over 50 percent of the sample of instructors agreed (either strongly or somewhat) with the statement, "Students who type usually get better grades." The instructors also tended to agree that typing helps students improve their spelling, punctuation, and organization.

The clear implication of this advertisement was that students will raise their grades if their papers are typed. A definite cause-and-effect relationship was suggested: Typed papers cause higher grades. If this message came through to high school and college students and their parents, we suspect that many people marched right out and bought a new typewriter, for surely the \$200 investment [which would be about \$600 in current day] would be worth it if the typewriter brought about better grades! **However, we're not convinced that the survey evidence supports this cause-and-effect inference. We believe an alternative hypothesis can explain why typed papers receive higher grades than papers that are not typed. Any ideas?**

.....

9. Moral Development across the Life Span

An emphasis on examining developmental processes across the life span led to many studies in cognitive as well as non-cognitive areas. One study by D. Bielby and D. Papalia examined development along a dimension that is linked to both cognitive processes and moral development as delineated by Lawrence Kohlberg's Stage Theory. Kohlberg's theory predicts that people progress through stages of moral development. Therefore, the researchers decided to see whether the moral development did increase with age.

The 72 middle-class participants in the study were sampled from six age groups (with six boys or men and six girls or women sampled from each age group). The age groups, I through VI, were 10-14, 15-19, 20-34, 35-49, 50-64, and over 65 years of age. Participants in poor health were excluded from the over-65 group.

Participants responded to stories presenting moral dilemmas. Because an age-by-gender analysis of variance showed no main effect of gender and no interaction of gender with age, subsequent analyses did not include gender as a factor. One of the results, based on a one-way analysis of variance of the six age groups, was significant.

Follow-up tests were then used to compare the specific mean levels of moral development for the six age groups. Significant differences were found between the age groups III ($X = 3.75$), IV ($X = 3.96$), and V ($X = 3.58$), and those four age groups differed from age group I ($X = 2.50$) and II ($X = 2.50$), although age groups I and II didn't differ from each other. Surprisingly, age group VI ($X = 2.92$) differed only from age group IV and age groups I and II.

The authors concluded that, for the most part, moral development increases with age (notwithstanding the oldest age). Do you see any limitations to the authors' conclusions?