

CRITIQUE OF RESEARCH ON ATTACHMENT THERAPY

BY RANDOLPH AND BY MYEROFF ET AL.

(BASED ON RANDOLPH'S SUMMARY,
ON-LINE AT www.attachcenter.org/research.htm)

Jean Mercer, Ph.D.

Jan. 6, 2001

1) Characteristics of well-designed research

The statement “garbage in, garbage out”, so often applied to computers, is equally applicable to empirical research. If the way in which a study is designed and carried out is inappropriate, the data collected can be little more than garbage, and the conclusions drawn from the data are unfortunately of the same nature.

Any casual observation can be called “research”, but the scientific community accepts the conclusions researchers draw from empirical studies only when the work follows certain design guidelines that have been developed over the last 100 years or so. Empirical research is thought to be well-designed and acceptable only when it has characteristics such as *falsifiability* and effective *isolation of variables*. This is especially the case when a researcher wants to understand whether one factor actually causes another, an essential question when we want to know whether a treatment is effective.

The term *falsifiability* involves the possibility that data could be interpreted to mean that a study's basic hypothesis or prediction is not correct when it is not, as well as that it is correct when it actually is. If a researcher has stated the problem so loosely that any data collected can be used to support the same hypothesis in some way, the design does not have falsifiability; no matter what the data look like, the researcher can still

claim, for example, that a treatment has been effective. The general public and some poorly-trained researchers think that their job is to “prove” a point by seeking out all confirming data and avoiding all disconfirming data, but such an approach lacks falsifiability and thus cannot be part of a well-designed study.

Isolation of variables involves, first, analysis of all the factors that play roles in a particular event, and, second, consideration of how each of them can be tested separately from all the others. When two factors are always tested together because the researcher does not realize that they can be separated, they are said to be confounded with each other. Confounded variables prevent the researcher from coming to an accurate conclusion about which of many factors causes which of many outcomes. For example, if all children receiving a treatment come from families that can pay for the treatment easily, and all children who fail to receive the treatment come from families who cannot afford the treatment, we cannot accurately conclude whether a difference between the groups was caused when the treatment helped certain children, or whether it resulted from the effects of economic stress on the families of the untreated children, or indeed what role might have been played by other variables confounded with these factors.

2) *How is falsifiability achieved?*

Falsifiability characterizes studies in which clear-cut outcome predictions are made before data are collected, and in which alternative explanations of data are not rejected without reason.

3) *How is isolation of variables achieved?*

Researchers try to achieve isolation of variables through three basic approaches, of different levels of acceptability.

a) The “gold standard” for isolation of variables involves an *experimental* design with *random assignment of individuals to a group that receives treatment or to a comparison group*. This is the design required by an American Psychological Association Task Force that made recommendations about claims that a treatment was effective. (The Task Force set a standard for evidence supporting a treatment, requiring publication of two independent studies, employing random assignment to groups, and demonstrating a statistically significant difference in outcome between a group who received the treatment and a comparison group who did not [Chambless & Hollon, 1998]). (Please note that a study is correctly called an experiment only when the researcher makes the decision about which treatment a particular individual will have. If the individual makes the choice himself or herself, or if the choice has been made by the family or some other factor, the study is not an experiment.)

Random assignment of individuals to groups means that the decision to treat a given person, or to withhold treatment, depends on an arbitrary factor such as the number appearing next in a random number table rather than on an existing characteristic of the person. If a person is assigned to a group on the basis of a personal characteristic such as ability to attend the treatment, the factor of treatment is not isolated from the personal characteristic, but confounded with it, so that it is impossible to conclude whether an outcome resulted from the treatment experience or from the personal characteristic.

When random assignment of individuals to groups is properly done, a relatively small number of participants can still yield a meaningful and statistically significant difference.

[Please note that random assignment to groups is not the same thing as random sampling, and that the latter is not an issue in this critique.]

b) A less effective attempt at matching of variables involves a *quasi-experimental* approach with *matching* of groups. In a quasi-experiment, the decision that a person will or will not receive a treatment is made in some way other than the researcher's choice. The individual may come into the study after already having had a treatment, or may decide to seek or avoid a treatment, or be placed for or withheld from treatment by a family decision involving finances or other family characteristics.

If the person or the family has made the decision about treatment, the variables of treatment and personal or family characteristics are confounded. Researchers may try to separate confounded variables, by a process known as matching, however. Matching is an attempt to make a treatment group and a comparison group as similar as possible in every way except their experience of the treatment. Matching tries to hold constant all factors that could influence the outcome except for the treatment itself.

The process of matching is complex and involves large numbers of possible participants, many of whom will be rejected from the study because they cannot be matched adequately with other participants. Even when matching is carefully done, it is never possible to conclude that all confounding variables were eliminated, and it is not permissible to conclude that the factor being studied (treatment vs. no treatment) was the actual cause of any differences in outcome.

c) A third approach to isolation of variables is *meta-analysis*. Meta-analysis is a complex statistical approach that deals with very large numbers of participants who took part in a variety of different but related studies. The assumption is that the personal or family characteristics of these people will cancel each other out, leaving the outcome to be determined only by what they all have in common (for example, the experience of a

treatment.)

4) Where does the attachment therapy research stand on falsifiability and on isolation of variables?

The hypothesis that children who receive attachment therapy will show a greater improvement in behavior problems than those who do not is clearly a falsifiable prediction; it would be possible for data to be interpreted to mean that there was no difference caused by the treatment. However, Randolph's interpretation of the data rejects alternative explanations of the outcome without good evidence for her decision. For example, noting that one study did not use a comparison group, she correctly comments that "It was possible that these same changes could have resulted from the passage of time, or from a difference in the child's living environment" but then goes on to claim that "given the fact [sic] that children with attachment disorder show an idiosyncratic resistance to improving... it was unlikely that the significant improvements found in the present study were due to something other than the therapy" (Randolph, p.9). She thus compromises the falsifiability of the study by an arbitrary rejection of alternative explanations.

As Randolph herself notes, her initial study has no comparison group; there is thus no isolation of variables possible. In discussing the Myeroff study, Randolph notes correctly that the design is quasi-experimental; the two groups were formed of children whose parents' decisions placed them either in the treatment or in the non-treatment comparison group. However, having said this, Randolph goes on to speak inappropriately of an "experimental" group and of a "control" group, neither of which forms part of a quasi-experiment.

The children who received treatment came from families who had applied to ACE and had brought their children for treatment. The non-treatment group was comprised of children whose families had applied for their treatment but “who, for some reason, could not work out a way to receive treatment” (Randolph, p.9). There was thus no isolation of the treatment factor from factors of family and/or personal characteristics of the children, and it is not possible to conclude that any differences were caused by treatment rather than by variables confounded with treatment.

Although no attempt to match the groups occurred, Randolph notes post hoc that the groups were similar in terms of gender, race, pre-adoption placements, age, and income of the adoptive parents. However, she gives no information about other family characteristics that could have a strong impact on a child’s behavior and behavior change. Some additional family characteristics that could have a strong effect include marital relationships and patterns of decision-making, the presence of other children with needs for unusual expenditures, and geographical location. It is notable that there is no evidence presented about age at adoption, the factor best known to affect the child’s development (Sharma, McGue, & Benson, 1996).

No meta-analysis has been possible in research on attachment therapy, because of the small numbers of studies available.

6) Were there other design issues that weaken these studies?

The issues of falsifiability and of isolation of variables are the major problems with the studies Randolph describes. However, some other points are of concern.

a) It seems curious that this research completely ignored the occurrence of *indiscriminate*

sociability among children with a history of poor attachment experiences, since this symptom is considered to be the major lasting difference between children adopted from orphanages and others (Zeanah, 2000).

b) The research Randolph describes also included a problem resulting from the nature of the Children's Behavior Check List (see <http://www.uvm.edu/~cbcl/>). This checklist involves ranking of children's characteristics relative to those of other children, and thus has the weakness that a given rater's definitions depend on his or her assumptions about "other children". This does not make much difference if we compare someone's ratings today with the same person's ratings tomorrow, but it makes it inadvisable to compare ratings by two different groups of raters, who may have quite different standards (see the work of Bartoshuk, discussed by Goode, 2001). (Randolph approaches this fact apparently unknowingly as she talks about differences between the ratings of foster parents and placing parents.)

c) It is difficult to tell from Randolph's presentation exactly how the statistical analysis was done. For example, analysis of variance is usually presented in the form of a summary table that shows the amount of variability as well as the means involved, but Randolph does not do this.

References

- Chambless, D., & Hollon, S. (1998). Defining empirically supported therapies. Journal of Consulting and Clinical Psychology, 66, 7-18.
- Goode, E. (2001, Jan. 2). Researcher challenges a host of psychological studies. New York Times, F1, F7.
- Sharma, A., McGue, M., & Benson, P. (1996). The emotional and behavioral adjustment of adopted adolescents: Part II. Age at adoption. Children and Youth Services

Review, 18, 101-114.

Zeanah, C. (2000). Research yields more signs of RAD in institutionalized children.
Brown University Child and Adolescent Behavior Letter, 16(12), 1, 6.