

# Overview of Research Designs for Rolfing® Structural Integration

by **Duffy Allen, M.S., Certified Advanced Rolfer, Rolf Movement Practitioner**

**A**t the most recent Annual Meeting, a group of interested individuals attended a series of meetings regarding research in the field of Rolfing. While the number of people in attendance (20 or so) may not seem overwhelming, it struck me that this was a significant percentage of Rolfers attending the larger conference. This indicates that there is a significant interest in establishing a pool of research from which to develop our professional credibility, to contribute to our theoretical understanding of Rolfing, and thereby, to refine our efficacy as Rolfers. As I listened to the discussion, I realized that there was a bias toward experimental research designs which require the most stringent controls, large numbers of subjects and highly artificial experimental environments. However, other types of research exist and contribute to theory development in clinical fields. Many of these methods are relatively easy to implement, even in our private practices, and require, at least in the initial stages, only a basic infrastructure and very little by way of financial resources. A discussion of the types of research common in diverse clinical settings could help broaden the discussion and draw in the interest and participation of more Rolfers.

## OVERVIEW OF THEORY FOR RESEARCH DESIGN

I will present here for the purposes of initiating a discussion, an overview of the broad categories of research that can be conducted in a clinical situation. In order to do this, I will present some general concepts about scientific research that will facilitate the dialogue. First, research can fall into two broad categories: descriptive or experimental.

Descriptive research includes a level of inquiry that includes a systematic recorded observation of a subject. An example of this type of research is a detailed anatomical dissection. When the findings of the dissection are recorded and documented, we can begin to explore the notions we have about how the body is organized structurally. Experimental research includes a number of research designs. Most familiar to us are those that require large subject pools and tight treatment controls. Single-subject experimental designs are commonly used as well, especially in evaluating clinical effects. These will be discussed later in this paper.

First, I would like to lay the groundwork for understanding research designs of all types.

## TECHNICAL CONSIDERATIONS

In order to compare different research designs, we need to establish some terms. The first concept we will explore is validity. Validity has to do with the ability of research to determine causal factors. A valid research design is defined as "one that presents evidence that a difference in performance across the two (experimental) periods is not due to extraneous conditions."<sup>1</sup> In addition, there are elements of an experimental design that contribute to clarity of the measurements to be made within the setting of a particular investigation. These fall into the subcategory of internal validity. This would include fundamentals like specific definitions for the purpose of the study. For example, a researcher may specify the location and degree of pain or range of motion or diagnosis of the subjects included in the study.

Additionally, it could include defining protocols or techniques or the guiding principles of the intervention. Strict operational definitions of terms bolsters the ability for accuracy of measurement or for multiple observers to measure a variable. "When a high degree of agreement is obtained between data collected simultaneously but independently by two or more observers....and/or by the same observer on two separate occasions...we can be more confident that our recordings are consistent."<sup>2</sup> This means that the ability to replicate studies is enhanced because the procedures may be accurately repeated by other investigators. The ability of a body of clinical investigators to reliably use clearly defined terms to describe an event contributes to the external validity, or generalizability, of a study which, in turn, allows it to contribute to the theoretical basis of the field. In other words, internal validity is a precursor to the generalization of the results to other cases.

In the most basic way, even a descriptive study, then, if properly executed, may contribute to our understanding of our work in a scientific way by generating a lexicon of operational definitions of terms. This is a primary stepping stone in developing a theory or a paradigm to guide a field of work. In philosophical terms, this level of investigation contributes to the theory core. "If a theory is a specification of a repertoire of mechanisms, its core is described when we specify the kinds of things it lists."<sup>3</sup> This means that building a theory happens in two steps. First, the elements that are critical to an event are extracted from extraneous variables. And next, the causes for the event are sought. The relevance of each particular element of the theory core is tested in accordance with the entire theory.

As is the case with all theories, a causal analysis of particular events, not classes of events, is critical to scientific testing. A theory core that derives the factors at work in an event broadens more traditional conceptions of theory. Statements which define the theoretical core make a sufficient contribution to theory in the absence of concrete evidence of causal mechanisms. Restrictions are then placed on the type and strength of the inferences which may be drawn from certain research designs based on the strength of their internal and external validity.

We have all heard the old axiom that correlation is not cause. A classic example of this in the realm of science is the concept of spontaneous generation. This was the pre-

vailing theory of reproduction of many types of creatures up until the end of the 17<sup>th</sup> century. It suggested that decaying matter would spontaneously generate flies, maggots, etc., rather than that these creatures had a phase of development that was not readily observable to the human eye that preceded the seeming eruption of these creatures fully formed. Likewise, even sophisticated modern research instruments and designs cannot demonstrate causality.

However, a very high correlation of events over time is generally accepted into a paradigm as a causal mechanism. Some research designs, because they more stringently control the variables at work in an event, can make stronger correlations between two events than those that have less rigorous control mechanisms. Remember that a theory seeks to define causal agents while a theory core seeks to define the basis of a theory in the absence of causal relationships. This means that different research designs contribute to theory development in different, yet important ways. It means that several levels of research design are important to understanding any field of work.

## RESEARCH DESIGNS

I have briefly alluded to different types of research designs used within the fields of biomedical and behavioral research. These can be divided into two broad categories, group designs and single subject designs. The following is by no means an exhaustive list of the types of research that are currently being conducted, but simply meant to give a contrasting overview of each and its relative strengths and weaknesses.

**Group designs.** Group experimental designs are generally associated with a deductive scientific philosophy. Specific results obtained from these designs contribute to inferences which appeal to general statements about behavior. That is, general statements about behavior are deduced from the specific examples gleaned from an experimental setting. Inferences drawn from group experimental designs garner support from the fact that the subjects within the group are heterogeneous, which in turn supports the capacity to draw generalizations from their behavior. Group designs compare the performance of one subject group against the performance of another. One group receives the experimental treatment while the other group receives no treatment and serves as the control. A statistical measure of central tendency (e.g., the mean or the median) of performance scores

for each group is computed. The difference between group performances is subjected to statistical analysis of significance. The treatment is credited with an effect on behavior when the two groups' performances are statistically different from each other.

These research designs are dependent on careful subject selection to ensure that each group is represented equally across such variables as age, sex, educational level, as well as qualities more specific to the question at hand. For instance in an investigation of low back pain, a researcher may want to pay special attention to the duration and severity of the pain as well as other factors. In this way, a "double blind" experiment does not mean that the subjects are selected totally randomly, but rather that they may fall within parameters deemed acceptable by the researcher or even matched on these variables across the control and experimental groups. The "blind" aspect alluded to above refers only to the knowledge of which individuals comprise a treatment group, which group is receiving the actual treatment and which is receiving a placebo.

The external validity of group designs is largely determined by their ability to be reproduced. Replication of investigations is important in determining the validity of inferences drawn from group experimental designs for several reasons. First, group designs are often administered in highly artificial environments. Replication of results across more naturalistic settings strengthens claims for treatment as a causal agent in changing performance. Second, statistical averaging across subjects in a group obscures the multitude of reactions to which a single treatment may contribute. Because performance is charted across a statistical central measure (e.g., mean or median scores, etc.) for potentially heterogeneous sample groups, no opportunity for delineating critical subject variables is provided. It is difficult to define the exact parameters of the elements of the correlation as they are blurred by statistical treatment. Other research designs address this issue more directly and will be discussed below. Third, statistical performance analysis limits comparisons to only one (technically, the null) hypothesis. "Within this context, testing only for the null hypothesis... encourages a disregard for all other individual observation made in the course of the experiment."<sup>4</sup>

And finally, group experimental designs traditionally do not incorporate true time series methodology. Performance is typi-

cally measure in pre- and post-treatment conditions. The absence of time series data detracts from the investigator's ability to draw causal inferences from obtained data. Of course, there are sophisticated time interactive designs which can mitigate these problems. However, they are complex to manage and require that the large groups of subjects remain in the experimental treatment for longer periods of time.

## SINGLE SUBJECT DESIGNS

While group research design gives us a snapshot view of a group of subjects, that is, a description at one particular point in time, other models take place over a period of time and give us the opportunity to document a subject over time and gives us information about the interaction of the subject with a stimulus (like a treatment or Rolfing sessions). Most of these longitudinal designs follow only a single subject and results are therefore constrained to less general application. Two general categories of single subject designs exist. I will outline each of these briefly here. Case study is a descriptive process and perhaps one of the most common means of relating a therapeutic process. Single subject experimental designs are more complex to describe, design and conduct, and so will only be outlined here.

Single subject experimental designs incorporate control mechanisms that contribute to the validity of inferences about causal mechanisms involved in a particular case. Single subject experimental research is generally conducted over a period of time. A baseline measure of a behavior or quality is taken over several sessions while no treatment is introduced. Then, the treatment or experimental variable is introduced. The behavior or quality which is being tested is measured over the course of the treatments. Again, this will happen over several sessions. The treatment will then be withheld (or changed, if the purpose is to compare two treatments) and behavior will be measured over several sessions. Then the experimental treatment will be reinstated as before for several sessions. In this way, behavioral patterns are demonstrated by the repeated measurements taken during the various experimental phases (treatment, baseline, withdrawal, reversal, etc.).

Two assumptions are required in evaluating the results of single subject experimental research. First, baseline behavioral measures of pretreatment behavior are assumed to be representative of future performance

if treatment is not introduced. And second, extraneous influences on performance are assumed to be equivalent across all experimental phases.<sup>5</sup> "Only repeated measurement over time can reveal behavioral patterns and changes in these patterns as treatment progresses."<sup>6</sup> The qualities being measured are graphically plotted to allow for visual inspection of the direction, magnitude and absolute level of change across experimental phases relative to control phases. Additionally, statistical regression lines may be plotted to demonstrate trends. Inferences can be drawn from these plots. When the quality being measured (dependent variable) changes at the time of an experimental phase change (independent variable, the treatment), the change is attributed to the treatment. The presence of control structures (baselines, withdrawal, or reversal of treatment) in these research designs increases the level of confidence in the dependence of behavior change on treatment.

External validity is obtained through replication of the change in behavior (or quality) within and across subjects. Within-subject replication is built in to most time series experimental designs. From there, replication of these investigations can lead to more general claims of treatment causality (that is, efficacy or theory development). Specific designs relevant to various clinical research questions (e.g., treatment effectiveness, comparative treatment effectiveness, and treatment component effectiveness) and procedures for their implementation can be found in various texts.<sup>7</sup> The current discussion will be limited to the design feature relevant to establishing correlated relationships (internal validity) and generalization (external validity) and thereby their contribution to theory core and theory. Generally, treatment conditions are maintained until a behavior stabilizes and is then periodically withdrawn or reversed in order to demonstrate control over the behavior fluctuations. Of course, in our field, this type of demonstration of control may lend itself to intolerable ethical or clinical effects (making the client worse). In this case, we might choose to focus on replicating treatment results across a number of individual subjects to demonstrate external validity.

### CASE STUDY

Another option in clinical investigation is to move to a descriptive research design in order to avoid these type of ethical dilemmas. A graceful case study falls into the spectrum of research designs which do not

readily allow researchers to make broad predictive statements, but instead are better suited to narrowing the parameters of an investigation. Specifically, we are looking for changes in the client that happen contrary to the expected or established trend. For example, if a client has had a consistent pain pattern for several months or years that is suddenly changed, we could assume a correlation to the treatment. In general, "a rapid and large change suggests that a particular intervention, rather than randomly occurring extraneous influences accounts for a pattern of results."<sup>8</sup> Of course, in a longitudinal study such as this, other factors (job promotion, relationship change, etc.) may be contributing, and therefore the correlation must be replicated with other subjects as well. A rule of thumb here is that the more diversity of the subjects for whom the treatment elicits change, the more powerful the treatment and the greater argument for a causal agent in the theory.

This process is dependent on having a number of practitioners documenting their work and for this data to be regularly made available to the professional body. In our field, we can begin to focus our research endeavors at all levels by beginning with implementing protocols for case studies and publishing the results in order to create a body of knowledge from which to devise more tightly controlled studies.

By following this process we move from a model that contributes primarily to theory core by helping to tease out the pertinent aspects of a treatment or subject to one that helps to develop the theory proper, by demonstrating correlations so strong as to be considered causal. This is well within the current paradigm operating in scientific philosophy in the behavioral sciences at the present time.

### CONCLUSION

This overview of research theory and design is presented in order to demonstrate a spectrum of investigative techniques that are available to us. Some of these require a good deal of experience in implementing and supervising and some can be used by each of us in the course of our private practices. Clinical evaluation strengthens our profession by clarifying the questions that we want to investigate as well as defining the terms within these questions. The upshot of this is that research is not only important to our professional development, a fact that has been pointed out previously in this journal, but also contributes to our

theoretical understanding of the work we do. My intention in laying out this overview is to initiate an interest in clinical research at all levels, but especially to demonstrate how the observations we make in our work with individual clients is important to our field. It is my hope that this discussion will lead to more research, and especially that we as practitioners see ourselves as a crucial resource in conducting and reporting the research we do in our offices by presenting information on single-subject designs. □

### NOTES

1. McReynolds, L.V. and Thompson, C.K., "Flexibility of Single-Subject Experimental Designs, Part I: Review of the Basics of Single-Subject Designs," *Journal of Speech and Hearing Disorders*, Vol. 51 (1986): 198.
2. Ibid: p. 201.
3. Miller, R.W., *Fact and Method: Explanation, Confirmation and Reality in the Natural and Social Sciences* (Princeton: Princeton University Press, 1987), pp. 142-143.
4. Barlow, D.H., Hayes, S.C., & Nelson, R.O., *The Scientist Practitioner: Research and Accountability in Clinical and Educational Settings* (New York: Pergamon, 1984), p. 57.
5. For more on this matter see, McReynolds, L.V. and Thompson, C.K., op. cit.; Ventry, I. and Schiavetti, N., *Evaluation of Research in Speech Pathology and Audiology* (Reading, MA: Addison-Wesley, 1986).
6. McReynolds, L.V. and Thompson, C.K., op. cit.; p. 200.
7. The ones with which I am familiar include: Barlow, D.H., Hayes, S.C., and Nelson, R.O., op. cit.; Kearns, K.P., "Flexibility of Single-Subject Experimental Designs, Part II: Design Selection and Arrangement of Experimental Phases," *Journal of Speech and Hearing Disorders*, 51 (1986): 204-214; McReynolds, L.V. and Thompson, C.K., op. cit.; p. 194-203; Ventry, I. and Schiavetti, N., op. cit.
8. Kazdin, A.E., "Drawing Valid Inferences from Case Studies," *Journal of Consulting and Clinical Psychology*, Vol. 49 (1981): 187.