

Evaluating research on the Feldenkrais Method from the outside. Some observations and suggestions.

By Jim Stephens, PhD, PT, CFP. FGNA Research Committee.

Abstract

As we debate the process and usefulness of research within our ranks, people are looking at us from the outside to try to evaluate the effectiveness of what we are doing in achieving the claims that we make. One of the ways we communicate our work is through the sharing of personal experience. Another way is the presentation of formal research that documents the outcomes of our work and suggests the context within which it may be most effective. This article will review some of the criteria that people use when they look at our work from the outside and discuss some of the conclusions about research on Feldenkrais Method based on those criteria. We have made a good start in addressing the outcomes of the work that we do but we have a long way to go to address the full range of the work that we do. Suggestions are made at the end for some next possible steps along the path of improving the research we do.

There is still a flourishing debate going on among practitioners as to whether it is possible, useful or even desirable to do research on Feldenkrais Method considering the strong individual and ideosyncratic nature of the work. Yet we talk about and point to changes in function as an outcome of the process we engage in. The challenge then for us is to find ways to capture the nature of these changes that we point to in ways that are valid and reliable. When we present ourselves to the outside world, we must be able to speak in ways that people understand and find meaningful. How does the world evaluate research on the Feldenkrais Method? What are the criteria? What are the conclusions about our work?

When national medical groups look at research and try to make recommendations for treatment based on what is in the literature they use several kinds of criteria. One kind looks at the type of design that was used as a measure of its usefulness. It ranks study designs by the reliability, validity and generalize-ability of the information which they can produce. Several design factors make good information more likely.

A control group provides baseline data, a group which can be used as a comparison against the experimental intervention. This kind of control group can eliminate spurious effects such as differences resulting from historical change, other normal experiences or normal processes of change within individuals. Random assignment of subjects to groups attempts to ensure that there is no bias expressed in selecting the groups and that demographic factors are balanced between the groups. Larger numbers of subjects, representing a wide range of the population in question provide greater generalize-ability.

Using these design criteria, studies can be stratified into five levels: Level I - a randomized controlled trial (RCT) with large numbers of subjects (more than 30) over long duration (months to years); Level II - a RCT with smaller numbers (fewer than 30 subjects) and/or shorter duration; Level III prospective cohort design with no baseline control. This is a design which has a control group but no pre-intervention baseline measures for comparison in either experimental or control groups. (This type of design is often used for surgical procedures where data can be collected only after the surgery is done). Level IV – cross-sectional controlled studies or retrospective cohort designs compare the performance of two different groups and assume that the differences noted are a result of some factor distinguishing the groups such as gender, age, race, training, etc; and Level V includes case studies of any size. The RCT is generally accepted to produce the most useful and valuable information. The case study method is generally considered the least useful or at least has a different kind of value when trying to extrapolate research results to a larger population. (The above criteria were used by the National MS Council in developing guidelines for treatment of fatigue in MS.)

Other kinds of research designs are commonly used. Single subject designs are similar to case studies but have more strict design criteria. Qualitative designs are used to map out the conceptual areas of a problem in early stages of understanding a research area. These are very good for capturing subjective responses and attitudes and developing a rich understanding of a person's experience but cannot be used to generalize to how others might respond. Survey research is used to collect descriptive data across groups of subjects for example, "What kinds of problems do clients present to Feldenkrais practitioners?". Each of these kinds of design has specific formal rules for how research is to be carried out.

Another kind of criteria used for evaluation of published research is whether a paper is peer reviewed. The peer review process assures that the paper has been evaluated, before publication, by people who are experts in the methodology and literature of that field. Peer review addresses such questions as: Do the subjects in the study meet specific entrance criteria? Are they balanced between groups by age, gender, task performance, etc. at the outset? Were they randomly assigned to groups? Were the measurement methods described in an understandable manner? Have the measurement methods been shown to be valid and reliable? Did the experimental design allow for other possible explanations of the results to be ruled out? Was the data analysis applied in a valid manner and was it appropriate for the design? Bottom line: peer review related to methods and results is important.

Clinical research is done in an imperfect world. Rarely does clinical research meet the highest standards of all of the above kinds of criteria. The MS Council, mentioned above, identified 700 studies related to the topic of fatigue in MS and found only 86 of them useful for developing clinical guidelines for management of fatigue in MS. There were no Level I studies available and only a few Level II studies which were drug studies. So members of the MS Council who are experts in the field had to base their development of clinical guidelines on consensus interpretations of lower level literature. We are not the only ones who do not have Level I literature to support our work.

How does research on the Feldenkrais Method stack up against these criteria? A recent article by Ives and Shelley (Work, 11: 75-90, 1998) provides an excellent review of research on the Feldenkrais Method up through 1996. They reported on a total of 42 research papers. Of these, 26 were non-peer reviewed, qualitative, descriptive case presentations covering a wide range of types of people and problems. These studies were criticized for their lack of methodological rigor. Protocols were insufficiently explained and standardized; there was no verification of information by triangulation, a method of comparing several different sources of information about the same question (e.g. computerized assessment of balance, assessment of different tasks requiring balance, and subjective report of balance performance). Other appropriate case study methodologies such as subject verification and approval of published information were not used. Moshe's work *The Case of Nora* was one of the reports reviewed here to which these criticisms apply. Ives and Shelley state that many of these reports make "extravagant" (p.85) claims which may fall into the category of speculation, which is not to say that they are not true, but methodologically unsound. In recent years, qualitative, case study methodology has become more rigorous and also much more respected if done following accepted guidelines.

Five of the reviewed studies were empirically based and nonpeer-reviewed. These studies used empirical measures rather than subjective interpretation, but were criticized for not testing or discussing the reliability and validity of the measurement tools, and for not having control groups. In the case of Shelhay (*Movement as a Model of Learning*, 1995) these criticisms may be unfounded, because only an abstract and not the full body of the work, which is published only in German, was reviewed. Six papers were peer-reviewed case studies. Half of these had serious flaws in methodology related to data collection and interpretation. The others were methodologically sound. These latter were papers by Narula (1993) and Schenkman (1989). (For references, refer to the research bibliography on the FGNA web site: <www.feldenkrais.com/research>.) Out of the total 42 studies reviewed only five were peer reviewed, randomized, controlled trials. All of these had methodological flaws which allowed their conclusions to be called into question. Among the 42 studies, none were Level I and only seven were Level II. Most of the rest were Level V, as described by the criteria of the National MS Council. Ives and Shelley conclude that the papers reviewed do not

present a convincing picture to support the anecdotal claims and that the empirical studies have not shown convincingly that "any of the positive findings can be directly attributable to Feldenkrais treatments." (p.85) They also conclude that "little evidence is provided that either acute or long term exposure to the Feldenkrais Method can promote changes that could not be obtained by using conventional treatments that may be simpler or more cost effective." (p.85) However they finally conclude that "The most support for the Feldenkrais Method comes not from any specific research findings but from the sheer number of reports that fit within a sound theoretical framework," and that more research on the Feldenkrais Method is warranted. (p.85)

The Ives and Shelley work does not include critiques of papers that were published after 1996. The research designs used and the quality of publications have improved since then. Twenty-nine research papers have been published in some form. Of these, 22 were published in peer-reviewed journals and 16 of these were randomized controlled-trial designs. Thirteen met Level I criteria, 4 met Level II criteria. There were 6 case studies. Most of these papers present evidence that suggests that when people engage in a process of Feldenkrais lessons in their various forms, there are measurable and significant changes (improvements) in function in both physical and psychological dimensions. Four papers illustrate this: 1) Laumer et al., 1997; 2) Bearman and Shafarman, 1999; 3) Lundblad et al., 1999 and 4) Stephens et. al. 2001. These are all methodologically sound presentations that report findings that are not as easily achieved by conventional treatments and that are either simpler or more cost effective.

Many practitioners are interested in doing research. There are four obstacles which limit us: money and other resources; skill in research design and methodology; a place of our own to publish this kind of research; and a large critically educated audience. The question arises: How can we continue to improve the quality and expand the sphere of research on the Feldenkrais Method? I would like to present several suggestions.

1. We should have a journal which publishes research on the Feldenkrais Method, which is peer reviewed by Feldenkrais practitioners and which meets rigorous methodological criteria. The question of peer review is interesting. We should be our own best peer reviewers. This journal should publish both quantitative and qualitative research
2. There must be training in research methodology not only for people who want to do research but also to raise the level of critical reading for practitioners who read literature on the Feldenkrais Method and in other areas. It should be noted that the best research done to date has been done mostly in association with degree programs at universities. We should have the skills within our ranks at this point to provide this kind of training within the training programs. Also there are now practitioners who are faculty in university programs which grant M.S. and PhD degrees who can provide training opportunities.
3. The training programs themselves then could begin to provide opportunities for conducting methodologically strong, clinically useful research. Frank Wildman, Osa Jackson and Mark Reese have raised this idea in the past. I challenge the TAB and trainers to take this need seriously and find a way to do it.
4. More money needs to be available for research. Over the past few years, \$2,000 to \$3,000 has been available annually from the FGNA. This has been used to support several small projects. There needs to be a larger commitment to research from the member practitioners of the Guild, who need to demand a larger budgeted amount, say \$15,000 to \$25,000 each year. I challenge practitioners to make this commitment and this demand of the FGNA, FEFNA and the IFF.
5. Now that the FGNA through FEFNA has 501 (c)(3) status it should be possible to find benefactors who will contribute larger sums of money to support research. This should become one of the primary functions of the Guild.