Proximal Change Experiments with Couples:
A Methodology for Empirically Building a Science of Effective Interventions for Changing Couples’ Interaction

John Gottman
Relationship Research Institute

Kimberly Ryan
New College of Florida

Catherine Swanson & Kristin Swanson
University of Washington


RUNNING HEAD: Proximal Change Experiments with Couples
Abstract

The concept proposed in this paper is a methodological innovation. The concept is that we can now do what might be called “proximal change experiments” with couples. The only objective of these experiments would be to improve the second of two conflict discussions, rather than having to change all aspects of the couple’s relationship for all time. One of the major objectives would be increasing the stability of positive affect during conflict discussions, a characteristic of the “masters” of marriage. With general systems concepts in mind, our nonlinear mathematical model for marital interaction and its parameters are described as well as the measurement network for evaluating proximal change interventions for distressed couples. The goal for positive affect becomes the addition of a positive “attractor” for the marital system. The results of a randomized clinical trial with five theoretically distinct brief interventions are then presented in terms of their ability to change the second of two conflict discussions. Differential treatment effects were obtained. Positive attractors were added when there were improving friendship and conflict management components in the intervention. The implications of these results are discussed.

Key Words: Marital, Intervention, Treatment, Positive Affect
Proximal Change Experiments with Couples:

A Methodology for Empirically Building a Science of Effective Interventions for Changing Couples’ Interaction

Re-Submitted 11/10/03

A series of prospective longitudinal studies of married couples was conducted in our laboratory. In these studies it was possible to predict with relatively high accuracy from observed communication variables, whether a couple eventually would stay married or divorce, and, if they stayed married whether they would wind up happily or unhappily married. This prediction has replicated over studies and could be based only on the observed interaction of the couple as they discussed an area of continuing disagreement in their relationship (Gottman, 1992, 1993, 1994, 1999), or an oral history interview about their relationship (Buehlman, Gottman, & Katz, 1992). The two major findings were that: (1) couples who later wound up stable and happy (even in very different types of marriages, Gottman, 1994) initially had a positive-to-negative-affect-ratio of 5.0 to 1 during a conflict discussion, whereas couples headed for divorce or unhappy stability initially had a ratio of 0.8 to 1; (2) some negative behaviors were more corrosive than others, particularly the “Four Horsemen of the Apocalypse,” namely, Criticism, Defensiveness, Contempt, and Stonewalling. Similar results have been obtained by other laboratories (e.g., Matthews, Wickrama, and Conger, 1996).

Subsequently, from categorical observational variables, we created a continuous summary variable, a “Dow-Jones-like Average” of a marital interaction, which was a time series for each partner that assessed the overall positivity minus negativity, at each of 150
six-second time blocks, of that person’s behavior. Using these time series, in two papers we were able to create a nonlinear mathematical model based on our predictions (summarized in Gottman, Murray, Swanson, Tyson, & Swanson, 2002). It was possible to create a mathematical equation for each partner, and within these two equations, there were parameters that formed a mathematical theory designed to explain our longitudinal predictions. Among these parameters were: (1) the “emotional interia” of each person, which measured the predictability of a person’s affect from that person’s immediate past behavior; (2) the “uninfluenced steady states,” which assessed how positive each person was independent of partner influence processes; (3) the “influence functions” that described how each person influenced their partner across the entire range of affect; (4) the “influenced steady states,” which assessed how positive each person was after influence processes occurred; (5) the “threshold” and “success” of each person’s repair attempts and the trigger threshold of negativity, which assessed how successful each person was at improving the communication once it passed a particular threshold of negativity, and, (6) the threshold and effectiveness at damping down positivity. It was therefore possible, based on our longitudinal predictions and the mathematical modeling, to derive parameters just from the couple’s observed interaction, which could be described as more “functional” or “dysfunctional.”

The idea presented in this article is that these parameters of the mathematical model make it possible to conduct proximal change experiments, whose only goal is to make the second of two conflict discussions look “better” than the first. By “better” we mean that the second interaction will look less like couples heading for divorce, or less like stable misery. These proximal change experiments can be very different from the typical complex multi-
component couples’ therapy outcome studies, which have a loftier goal of dramatically changing every aspect of the entire marriage forever. The proximal change experiments have a more modest goal. They can evaluate the success of very specific, even brief, and theoretically distinct interventions.

Why is this an advantage? One problem with a successful complex, multi-component intervention is that it is hard to identify the active ingredient of change. Therefore, it is hard to systematically build a science of change interventions. Proximal interventions can be more specific, smaller and more modest, since the goal of the proximal intervention is a smaller goal. Across many proximal experiments, together with the results of therapy intervention studies, results could be used to assist in building a library of change interventions, each of which accomplishes some specific goals, but perhaps fails to accomplish other goals. Proposing a methodology that would assist in building that library of interventions is the objective of this article.

Hope in developing such a methodology to assist clinical intervention and prevention contrasts with a broader sociological perspective (perhaps through affecting social policy) that seeks to alter socio-cultural factors related to marital disruption and discord, such as age at marriage, pre-marital pregnancy, economic distress, and so on. Our view needs to be seen as a clearly limited perspective that suggests that part of what may be necessary in marital success is to improve interactive behavior.

In this paper we illustrate this proximal change experiment approach by evaluating five specific interventions. One intervention (called IMPROVE FRIENDSHIP) was a brief one-day workshop designed only to increase the quality of friendship and intimacy in the relationship. Another brief intervention (called MANAGE CONFLICT) was designed to
decrease dysfunctional conflict and increase functional conflict processes. Another brief intervention (called BOTH) was designed to increase the quality of friendship and intimacy in the relationship and to decrease dysfunctional conflict and increase functional conflict processes. Another, somewhat longer intervention was designed to do BOTH, plus an added nine sessions of marital therapy that also focused on relapse prevention (the intervention will be called BOTH+THERAPY; a manual for the therapy is in Ryan, 2000). Interestingly, these nine sessions of therapy began after the post-assessment interaction session, so that at the time of post-assessment this group of couples actually had exactly the same intervention as the BOTH group, except that they also had only the additional expectation of receiving therapy after the workshop. Hence, for the purposes of this article, the BOTH and the BOTH+THERAPY groups differed only in the expectation of receiving more therapy. The final intervention was a Bibliotherapy intervention (called BIBLIOTHERAPY). This was a very low-cost intervention, every couple in the bibliotherapy group received Gottman and Silver’s (1999) book The seven principles for making marriage work, and up to three hours of telephone consultation with a doctoral student in clinical psychology.

**Objectives of the Interventions.** We can derive the objectives of the interventions from our longitudinal research, examining two questions.

**Reducing negativity.** The first question is, what did our longitudinal research find that distinguished (at Time-1) the happy couples from the unhappy ones? The second question is, over time what predicts divorce? The answers to these two questions in terms of the mathematical model parameters led to three goals of our interventions, namely that (1) the uninfluenced steady state, and the (2) influenced steady states should become less negative,
(3) the onset of repair attempts should begin at a less negative threshold. These results are related to negative affect and its escalation in the couple’s conflict discussion.

**Increasing positivity.** Our other findings were about positive affect. In our longitudinal study of newlyweds we found that positive affect during a conflict discussion was the single best predictor of both stability and happiness over a 6-year period (Gottman, Coan, Carrere, & Swanson, 1998). We have noticed for many years that the couples we have called the “masters of marriage,” defined as couples who stay married both with reasonably high marital satisfaction, seem to have a shared sense of humor and can be affectionate even when they are discussing an area of continuing disagreement. This positive affect acts in the service of de-escalating the conflict and maintaining physiological calm. These results held for newlyweds, and in a longitudinal study of middle-aged and senior couples in first marriages, we also found that humor and affection was a characteristic of happily married stable older couples (Carstensen, Levenson, & Gottman, 1995).

Although this information about positive affect is interesting, it is very difficult to apply clinically. How is one to induce positive affect into a conflict discussion? It appears to be ineffective to simply admonish distressed couples to be more positive when they are conflicting. Vincent and Friedman (1979) tried to do this in a fascinating experiment. They simply asked unhappily married couples to fake good for the camera, by pretending that they were happily married. They found that distressed couples could not do this successfully. When asked to act happily married, even blissful, during an argument, couples could not hide their non-verbal behavior. Negative affect leaked out during the conflict discussions, so that trained observers could easily tell which couples were distressed and non-distressed simply by their non-verbal cues.
Hence, simply asking a couple to pretend to be positive during conflict did not create a more positive interaction. Therefore, a fourth goal of our interventions was: (4) to increase the occurrence of positive-positive stable steady states (attractors) in the marital conflict discussion. This goal is far different, and far more difficult to achieve, than decreasing negativity. It is also more ambitious than simply increasing positivity. A positive attractor is a stable steady state that repeatedly draws a couple toward this positive place in phase space.

In addition to these four objectives, theoretically, we would also expect that our interventions should: (5) decrease emotional inertia, (6) increase the influence of positive affect (measured by the slope of the influence function in the positive affect ranges), (7) decrease the influence of negative affect (measured by the slope of the influence function in the negative affect ranges), and, (8) increase the effectiveness of repair attempts. For these latter, theoretically derived goals at this time there is no supporting empirical evidence.

**Our Mathematical Model of Marital Interaction**

History of our modeling efforts. The application of applied mathematics to the study of marriage was presaged by Von Bertalanffy (1968), who wrote a classic and highly influential book called *General System Theory*. This book was an attempt to view biological and other complex organizational units across a wide variety of sciences in terms of the interaction of these units. This book had a large impact on the family general systems theorists, although the vision of a mathematical basis was lost after von Bertalanffy. Our mathematical modeling work is a return to von Bertalanffy’s original dream. Specifically, von Bertalanffy’s dream was that the interaction of complex systems with many units could be characterized by a set of values that changed over time, denoted $Q_1, Q_2, Q_3$, and so on.
We can presume that each Q variable indexed a particular unit in the “system,” such as mother, father, and child, and, furthermore that these variables measured some relevant characteristic of a person that changes over time, such as the number of angry facial expressions per unit time. Actually, the Q’s were quantitative variables that von Bertalanffy never specified. However, he thought that the system could be best described by a set of ordinary differential equations of the form:

\[
\frac{dQ_1}{dt} = f_1(Q_1, Q_2, Q_3, \ldots) \\
\frac{dQ_2}{dt} = f_2(Q_1, Q_2, Q_3, \ldots) \\
\text{and so on.}
\]

The terms on the left of the equal sign are time derivatives, that is, rates of change of the quantitative sets of values \(Q_1, Q_2, Q_3,\) and so on. The terms on the right of the equal sign are functions, \(f_1, f_2, \ldots,\) of the Q’s. Von Bertalanffy thought that these functions, the \(f\)’s, would generally be nonlinear. The equations he selected have a particular form, called “autonomous,” meaning that the \(f\)’s have no explicit function of time in them, except through the Q’s, which are functions of time. These are, in fact, the type of equations we worked with. However, von Bertalanffy presented a table in which these nonlinear equations were classified as “Impossible” (von Bertalanffy, 1968, p.20). He was referring to a very popular mathematical method of approximating nonlinear functions with a linear approximation.

However, it was not the case is that these nonlinear systems of equations were “impossible” even at the time of his writing; unfortunately, von Bertalanffy was unaware of the extensive mathematical work beginning in the 19th century with Poincaré (Poincaré, 1993) on nonlinear differential equations, chaos, and fractal theory, which was only to
become widely known in the 1980s. In fact, in recent times the modeling of complex deterministic (and stochastic) systems with a set of nonlinear difference or differential equations has become a very productive enterprise across a wide set of phenomena, across a wide range of sciences, including the biological sciences.

We thus applied a relatively old approach to the new problem of modeling social interaction using the mathematics of nonlinear difference and differential equations. These equations express, in mathematical form, a proposed mechanism of change over time. The equations do not represent a statistical approach to modeling, but rather they are designed to suggest a precise theoretical mechanism of change. This method has been employed with great success in the physical and the biological sciences (e.g., see Murray, 1989). It is a quantitative approach that requires the modeler to be able to write down, in mathematical form, on the basis of some theory, the causes of change in the dependent variables. For example, in the classic predator-prey problem, one writes down that the rate of change in the population densities is some function of the current population densities. The equations are designed to write down the precise form of rates of change over time.

The ideal mathematical technique for describing change is the area of differential and difference equations. Mathematicians often used linear terms or linear approximations of nonlinear terms, and they usually gave very good results. In fact, most of the statistics psychology uses are based upon linear models. However, in the area of nonlinear differential equations, when linear equations are used as estimates (which simply assume that rates of change follow generalized straight-line functions of the variables rather than curved line functions), the estimated linear models were generally unstable, except very close to the system’s steady states, or “attractors” (see, for example, Strogatz, 1994, particularly his
discussion of Romeo and Juliet’s love with a linear model). This instability of linear models as one moves away (often even slightly) from the steady states of the system was a serious problem.

In recent years it has also become clear that most systems are complex and must be described by nonlinear terms. At first, this seemed like a real drawback, because it is usually not possible to obtain solutions in closed mathematical form with nonlinear equations. Interestingly, however, it has turned out that by employing nonlinear terms in the equations of change some very complex processes can be represented with very few parameters. Because solutions usually could not be obtained in closed mathematical form, graphical methods had to be developed for understanding the character of solutions to these equations. For this reason the methods have been called “qualitative,” and visual graphical methods are central in the mathematics. For this purpose, numerical and graphical methods have been developed such as “phase space plots.”

While these qualitative methods initially may seem like a drawback of nonlinear models, we suggest that these visual approaches to nonlinear mathematical modeling can be very appealing because they can engage the intuition of a scientist working in a field that has no mathematically stated theory. If the scientist has an intuitive familiarity with the data of the field, our approach may suggest a way of building theory using mathematics in an initially qualitative manner. Therefore, the theoretical decision to complete the mathematical work that von Bertalanffy suggested was motivated by the goal of creating a qualitative mathematical language for describing social interaction that engaged the scientist’s intuition for the purpose of building theory.
The attractors in a marital interaction. Once we write down the equations of marital interaction, the first question mathematicians ask is “Toward what values is the system drawn?” To answer the question we define a “steady state” as one for which the derivatives (on the left side of the Bertalanffy equations) is zero. This means that the system at a steady state does not change. If these steady states are stable, then, they are called “attractors” of the system. “Stability” means that if the system is perturbed away from these states, it will be drawn back to them, very much like gravitational attractors draw mass back toward the attractor. However, some steady states are unstable. Instability means that if the system is perturbed away from these states, it will be drawn away or repelled from them. The form of the equation determines whether a steady state is stable or unstable and, if it is stable, the strength of the attraction.

The goal of our modeling was to dismantle our two Dow-Jones-like time series (one for each partner) into components that had some theoretical meaning. Instead of differential equations, we used difference equations. We wrote the wife’s time series at time $t+1$ as $W_{t+1}$ and the husband’s time series at time $t+1$ as $H_{t+1}$. Then we broke the time series into a sum of the following components: (1) a constant that would indicate where each person started the interaction, before influence processes began; (2) an autocorrelation term that described how predictable each person was from his or her immediate past ($W_t$ or $H_t$); (3) a cross-correlation influence function that described the influence of the partner. Hence, our equations were:

$$W_{t+1} = a + r_1 W_t + I_{HW}(H_t) \quad (EQ \ 1)$$

$$H_{t+1} = b + r_2 H_t + I_{WH}(W_{t+1}) \quad (EQ \ 2)$$
The constants are $a$ and $b$ assess the “initial” affective state of husband and wife before influence processes have had their effects. The $r_1$ and $r_2$ are the autocorrelations, which we call “emotional inertia.” Emotional inertia describes how influenced each person is by his or her immediately past affect. High emotional inertial limits how influential the partner can be (and conversely). We have learned (see Gottman, Murray, Swanson, Tyson, & Swanson, 2002) that the “uninfluenced” parameters are actually a function of the immediate past history of the couple’s interaction as well as enduring personality characteristics of that partner.

**Building theory.** The theoretical portion of writing our equations lies in writing down the mathematical form of the influence functions (the $I_{HW}(H_t)$ and $I_{WH}(W_{t+1})$). An influence function is used to describe the couple’s entire interaction. The mathematical form is represented graphically, with the x-axis as the range of values of the Dow-Jones dependent variable (positive minus negative over a 6-second time block) for one spouse and the y-axis the average value of the dependent variable for the other spouse’s immediately following behavior, averaged across time or turns at speech. This latter point is critical, and it may be unfamiliar to social scientists: The influence functions represent averages across the whole interaction.

When we began modeling, the mathematicians we worked with (James Murray and his students) asked us to give them information so that they could write down the functional form of what determined rates of change of each person’s behavior. We were initially at a loss. However, once we began asking the question, “What did we know about marital interaction that could help us write down the mathematical form of the influence functions?,” we decided that one consistent result that had been obtained by many
laboratories studying marital conflict interaction with observational methods was that negative affect was a better correlate of marital satisfaction and predictor of longitudinal course than positive affect. We thought that perhaps this means that we should expect that the theoretical form of the influence functions would probably be bi-linear, with a steeper slope in the negative affect ranges than in the positive affect ranges. Thus, we would expect that the influence functions would have to be somewhat asymmetric, or nonlinear. Each influence function therefore had two slopes in the two regions are the important parameters for the bi-linear form of the influence functions (see Figure 1). It may be useful to explain that Figure 1 is a theoretical model of influence from husband to wife, summarizing the entire interaction over time. It provides a picture of the average influence that the husband had on his wife’s next data point across the entire range of his affect. There is a similar influence function from wife to husband. Many other types of influence functions could have been proposed (and we did experiment with some others). In the computational procedure, the theoretical form is then used to estimate the influence functions.

**Estimation Procedure.** The estimation procedure is to define a subset of observations for which we can safely assume that the influence function is zero. We took both partners being zero (neutral affect) as those values. Then our equations reduce to:

\[
W_{t+1} = a + r_1 W_t \quad (EQ \ 3)
\]

\[
H_{t+1} = b + r_2 H_t \quad (EQ \ 4)
\]

And we can estimate \(a, b, r_1,\) and \(r_2\) by least squares fitting for the subset of points. We then return to the full data set, and compute the influence functions from the data as:

\[
W_{t+1} - \left[ a + r_1 W_t \right] = I_{HW}(H_t) \quad (EQ \ 5)
\]

\[
H_{t+1} - \left[ b + r_2 H_t \right] = I_{WH}(W_{t+1}) \quad (EQ \ 6)
\]
In this estimation procedure (Cook et al., 1995; Gottman, Swanson, & Murray, 1999; see especially Gottman, Murray, Swanson, Tyson, & Swanson, 2002), we were able to estimate all of the parameters in the equations, and some important derived parameters.

**Modifying the Influence Functions.** We later added to this bilinear graphic idea of the influence functions a “repair” term that would, after a particular threshold, raise a negative graph with some effectiveness. The repair term down-regulates negativity. We also later added to this graphic idea a “damping” term that would, after a particular threshold, decrease a positive graph with some effectiveness. The damping term down-regulates positivity. We initially had little idea what damping might mean, but thought that it would be harmful. However, as we began working with stability regions in phase space, we learned that damping would at times create a positive stable state when none might have existed, so we concluded that damping can be a useful part of marital interaction. The important point is that the form of the influence function is our “theory” about relationships. Note that this means we can change our theory if it is not working to adequately represent reality (i.e., our data). In fact we have experimented with alternative forms of the influence functions. To get some idea of how theory could enter the design of the influence functions, in the bilinear function there were two slopes, the influence of positive and negative affect ranges on the partner. Note that in Figure 1 we have drawn the slope for negativity as steeper than the slope for positivity, and we have done so because that is consistent with marital process research. Typically, in the context of resolving a conflict among married couples, negative affect has a bigger impact in correlating with marital satisfaction and in the prediction of marital outcomes than does positive affect. We call this phenomenon the “triumph of negative over positive affect.” We discovered that for married couples the slope of the
influence function in the negative domain is larger than that in the negative domain, \( a_i < b_i \).

As we have noted, we called this the “triumph of negative over positive affect,” which refers to the general finding in much observational research of conflict discussions in married couples that negative affect is a better predictor of outcomes (happiness, stability) than positive affect. There is some reason to believe that during courtship these slopes may be reversed, that there is the triumph of positive over negative affect, that is, \( a_i > b_i \). One may tend to minimize one’s partner’s negativity and emphasize one’s partner’s positivity, a state that has been called “limerance” during courtship (Noller, 1996). We also found some evidence that for gay and lesbian couples there is the triumph of positive over negative affect, perhaps because there are fewer barriers to leaving than in marriages, and courtship must be continually renewed. In this way the model is informed by previous empirical research. There are, of course, two influence functions, one for the husband and one for the wife. Differences in slopes were our initial operational definition of power imbalances in the relationship, which is defined separately for positive and negative affect.

The central parameters we wound up with are: (1) thresholds of positivity and negativity. These thresholds represent the values of how negative one partner’s negativity has to get before it starts having an impact on the other person, and how positive the partner’s behavior has to get before it starts having a positive impact on the other person; (2) the emotional inertia (positive or negative) of each spouse, which is each person’s tendency to remain in the same state for a period of time. The greater the emotional inertia, the more likely the
person is to stay in the same state for a longer period of time, and less likely to be influenced by the partner; (3) A derived parameter from knowing both this starting value and the emotional inertias of both people (one that emerged from solving the equations) was the couple’s uninfluenced steady state, which is their average level when their spouse did not influence them (the influence function was zero). (4) Another derived parameter was the influenced steady state of the interaction, which is a steady state of the system, after influence has taken place. We would expect that the influenced steady state is more positive than the uninfluenced steady state in marriages that are stable and happy; that is, the influenced steady state answers the question, To what extent does the marital interaction pull each individual in a more positive or a more negative direction compared to the uninfluenced steady state? (5) The shape of the influence function, given by the following parameters: The linear slope in the positive affect region, the linear slope in the negative affect region, and the repair and damping terms (threshold and effectiveness).

Determining the attractors. The form of the influence function determines the steady states of the system of equations. It is of considerable importance to find the steady states of the phase plane, the influenced set points, and this is accomplished by finding those points were the null clines intersect. Remember that the null clines are determined by the equations. Finding the null clines is accomplished mathematically by plotting them. Null clines involve searching for steady states in the phase plane; they are theoretical curves where things stay the same over time. A person’s null-cline is a function of their partner’s last score and gives the value of their own score when this is unchanged over one iteration. Mathematically, this is written as:

\[ W(t+1) = W(t) = W. \]
This last equation says that, for the wife’s behavior, things stay the same over time, and that is precisely how we find the shapes of the null clines. Plotting null-clines and finding their intersections provides a graphical means of determining the steady states. First we begin with simple algebra in which we substitute $W$ for all the wife terms. This process gives us:

\[ W = r_1 W + a + I_{HW}(H_t), \]
\[ W - r_1 W = I_{HW}(H_t) + a, \]
\[ (1 - r_1) W = I_{HW}(H_t) + a, \]
\[ W = \frac{I_{HW}(H_t) + a}{1 - r_1}, \]

That last equation is the wife’s null cline. It’s the curve where she doesn’t change. When we do the same analysis for the husband’s null cline, and recall that the steady states are the intersection of the null clines, this then gives the final form of our null-clines as:

\[ W(H_t) = \frac{I_{HW}(H_t) + a}{1 - r_1} \]
\[ H(W_{t+1}) = \frac{I_{WH}(W_{t+1}) + b}{1 - r_2} \] (EQ 7)

Therefore, we have discovered by simple algebra that our null clines are simply the influence functions, scaled (by $1 - r_1$, or $1 - r_2$) and moved (by $a$, or $b$). In other words, we have shown that the null clines have the same shape as the influence functions, they are moved over (translated) by a constant, and they are scaled by another constant. Null-clines play an important role in mathematical analysis since they give a visual indication of the dynamics of the system.

As we noted, the attractors, equilibria, or steady states are determined by examining intersections of the null-clines, since, by definition, if the system started at this point then it would stay there. The stability or instability of these steady states to perturbations can also be computed mathematically (determined from the eigenvalues of the system of equations,
see Gottman et al., 2002). Since we have not specified the functional form of the influence functions, we can only proceed qualitatively. To derive the influenced steady states of a marital system, the pair of equations (EQ. 7) can be solved graphically. Therefore, if we plot the two curves from EQ. 7 their solution will be given by any points where the curves intersect. They are given by Figure 1 based upon what we know about couples’ interaction to generate a functional form for the influence functions, and we now turn toward a consideration of what the influence function slopes mean psychologically.

Figure 2 is a plot of the two steady states of a marital system with the bilinear influence function theory. In nonlinear dynamic mathematical modeling (e.g., see Murray, 2002) the form of the influence functions determine what are called the “null clines” of the equations, which are the curves in the x-y (e.g., Husband-Wife) phase space where the wife is steady and the husband is also steady. In the differential equations form of the model these null clines would be obtained by setting time derivatives to zero. Where the null clines intersect determine the influenced “steady states” of the system of equations. Steady states can be either stable or unstable. If an influenced steady state is stable, the influenced steady state is also called an “attractor” of the dyadic system, or a homeostatic set point. There may be more than stable attractor one for any dyad. These attractors may be compared to gravitational attractors, and we can estimate their strength. We also compute a diagram of the vector field in phase space determined by the equations, which give flow diagrams of likely trajectories for a couple in phase space. These vector fields tell us how the system is likely to behave for any starting point.
What have we gained with all this mathematics? What is the advantage of the mathematical modeling over just summarizing the interaction with variables that describe how much positivity and negativity there was in the interaction? The advantage is that we now have a mathematically-based language for describing central processes of the interaction in general systems terms, just as von Bertalanffy imagined. We can describe how the interaction begins (the uninfluenced steady states), how much emotional inertia each person has, how each person influences the partner very specifically with the influence functions, and the ultimate result of the influence, the influenced steady states. There is now theoretical causal language about interaction processes, language that is mathematical, and which can be altered as needed to fit the data.

Methods

Participants

Beginning with a volunteer sample of 400 couples from the Greater Seattle Metropolitan area recruited with radio and television interviews, newspaper advertisements, and flyers, one hundred couples were selected for participation so that a volunteer sample could be recruited that was representative of the demographics of the city of Seattle. Demographic information and telephone Locke-Wallace marital satisfaction scores were first obtained from all 400 couples. They answered a set of telephone survey questions that assessed their availability to attend the intervention, demographic characteristics, marital status, years
married, marital satisfaction, and health. To be eligible for the study, couples were required to be living together, to be legally married, to be able to attend the scheduled marital intervention and at least one spouse had to have a marital satisfaction score below 93 on the Locke-Wallace Marital Adjustment Test (½ standard deviation below the mean on this measure; Locke & Wallace, 1959). If the couples met the initial selection criteria, they were mailed a pre-treatment packet of questionnaires to complete individually, a questionnaire consent form, a cover letter with directions, and a prepaid return envelope. Following the application of the screening criteria (described below), 100 distressed couples were then selected to match the racial demographics of the Greater Seattle Metropolitan Area (based on the current City of Seattle’s Planning Commission Report), and these demographics were balanced across the five groups of the study.

**Procedure**

*Experimental Design.* Using a stratified random sampling method to insure that all five intervention groups were balanced by race and ethnicity, all couples passing through the screening process were randomly assigned to one of five treatment conditions. The couples were then scheduled for pre-workshop and post-workshop laboratory sessions.

*Intervention.* The interventions were administered in a psycho-educational workshop format, with lectures, demonstrations and exercises completed by husband and wife. These workshops were conducted by Drs. John and Julie Gottman at the University of Washington. There is evidence to suggest that a purely psychoeducational format can be effective in treating marital problems (Kaiser, Hahlweg, Fehm-Wolfsdorf, & Groth, 1998). All of our exercises are given in Gottman (1999) and in Gottman & Silver (1999). The interventions were: (1) IMPROVE FRIENDSHIP: a component with lectures and exercises
designed to enhance a couple’s friendship (building love maps, building fondness and admiration, and emotional connection through turning toward one another in everyday interaction); this intervention also contained an abbreviated section on repair by processing fights by having a recovery conversation after the fight, using our Aftermath of a Fight procedure, and, (2) MANAGE CONFLICT: a component with lectures and exercises designed to teach a couple to regulate conflict through (a) the management of gridlocked perpetual conflict (moving from gridlock to dialogue), with our dreams-within-conflict exercises, and (b) the management of solvable conflicts through softened startup, accepting influence, effective repair, physiological self and partner soothing, taking effective breaks, and compromise; (3) BOTH: a two-day workshop that had as its first day IMPROVE FRIENDSHIP and as its second day MANAGE CONFLICT; (4) BOTH+THERAPY: followed the same procedures as the BOTH condition, with nine sessions of marital therapy, which began after the post-workshop assessment; and, (5) a BIOBLOIOTHERAPY condition in which each couples received a copy of Gottman & Silver (1999). There were three different manuals for these three workshops distributed to couples and followed in the workshops. Thus, some couples received the full two-day workshop (improving friendship and regulating destructive marital conflict) and others one of the more limited one-day workshops, focused on either improving the couples’ friendship or in regulating destructive marital conflict. The manual for the one-day friendship enhancement workshop was not quite a carbon copy of the first day of the two-day workshop. For the sake of providing a reasonably complete theory of how marriages function, exercises were added that discussed the repair of negative interaction and processing a fight, and lectures were added that claimed that the enhancement of friendship was all that was necessary for repair to be
effective. The manual for the one-day conflict regulation workshop was a carbon copy of Day 2 of the two-day workshop. Each day of the workshop lasted 8 hours.

**Marital procedure for obtaining observational data.** Observational data of conflictual marital interaction were obtained to directly examine interactive marital behavior. In the laboratory couples were asked to complete the Couple’s Problem Inventory (Gottman, Markman, & Notarius, 1977), which measures the severity of various marital problems. Items include standard marital problem areas such as in-laws, finances, and sex. Each item was rated on a scale from 0 to 100, with higher scores signifying that the problem is considered more severe. The facilitator then reviewed the results of this questionnaire with couples to reflect on the issues they rated as most problematic, and helped them to choose an issue to use as the bases for a discussion of disagreement. This process of interviewing the couple about their areas of disagreement helps to insure that the couple has identified a good clear, current and emotional area to discuss. After choosing topics for the discussion, couples were asked to discuss their chosen topics for 15 minutes. Using two cameras that each obtained a full-face view of each couple, merged in a split screen, all videotapes were coded by two independent observers.

**Behavioral coding of the marital interaction.** The Specific Affect Coding System (SPAFF; Gottman, McCoy, and Coan, 1996) was used to code the couples' conflict interactions. The system was used to index specific affects expressed during the session of marital problem resolution. SPAFF focuses solely on the affects expressed. The system draws on facial expression (based on Ekman and Friesen’s Facial Action Coding System; Ekman & Friesen, 1978), vocal tone, and speech content to characterize the emotions displayed. Coders categorized the affects displayed using five positive codes (interest, validation, affection,
humor, joy), 10 negative affect codes (disgust, contempt, belligerence, domineering, anger, fear/tension, defensiveness, whining, sadness, stonewalling), and a neutral affect code.

**Weighting of the SPAFF Codes.** For the mathematical modeling we use a weighting scheme derived from previous prediction research (Gottman, 1994). A numerical value is calculated for the SPAFF codes for each 6-second time block separately for each partner by taking the sum of positive codes minus the negative codes using the following weights: Disgust = -3, Contempt = -4, Belligerence = -2, Domineering = -1, Anger = -1, Fear = 0, Defensiveness = -2, Whining = -1, Sadness = -1, Stonewalling = -2, Neutral = 0.1, Interest = +2, Validation = +4, Affection = +4, Humor = +4, and Excitement/Joy = +4. This weighting yields a potential score range of -24 to +24. For each couple this created two time series, one for the husband and one for the wife, each with 150 data points, one series for the “initiator” and one for the “partner”. The inter-observer correlations for these weighted data for married couples in one study was .90 (p<.001) during the conflict discussion. For more data on the inter-observer correlations obtained in other studies with the SPAFF, see Gottman (1994b).

**Analysis plan for assessing the effectiveness of the interventions, and differential treatment effects**

For each of the mathematical parameters, a repeated measures t-test was conducted to determine whether that parameter had changed following intervention. Because the interventions were so different theoretically, we can examine what deficits each of the interventions left in its wake, and so draw conclusions about the role of these processes in effective conflict discussions.
Results

Preliminary Analyses

Demographics. Chi-square analyses and t-tests revealed no significant differences between any of the groups on any of the demographic variables (ethnicity, years married, age, income). As such, the demographics of the groups are reported together. Consistent with the city of Seattle’s demographics, the sample was predominately White or European American, with 68% of both the husbands and wives identifying as such. The remainder of the sample was split fairly evenly among African American, Asian American, and Hispanic participants. Multiple ethnic/racial identities were permitted in our demographics. For husbands and wives, respectively, the breakdown was Hispanic-American, 6.5%, 8.7%; African-American 9.7%, 10.9%; Asian-American, 4.3%, 8.8%; Pacific Islander or Hawaiian, 3.2%, 2.1%; Native American, 3.2%, 1.1%. The mean number of years the participants had been married to their current partners was 13.0. Wives in the study had a mean age of 41.99 (SD=11.41) years, while husbands had a mean age of 44.64 (SD=12.19) years. The average income for wives was $20,800 and for husbands was $48,900. The screening mean marital satisfaction scores were: for husbands 80.77 (SD=23.49), and for wives 74.85 (SD = 22.45). This sample was more distressed than the typical marital therapy study (for example, Greenberg & Johnson, 1988 reported two studies; in the first study the pre-intervention marital satisfaction levels for the experimental group were 92.8, and 86.3 in the second study). Couples were effectively randomized within strata across groups so that each treatment group was equivalent by race/ethnicity and income.
How the interventions changed the mathematical model parameters

**Emotional inertia.** Inertia is a measure of the predictability of a person’s affect from that same person’s immediate past behavior. It limits how much influence the partner can have. The more emotional inertia a person has, the less capable they are of being influenced by their partner. In our longitudinal research, high emotional inertia was predictive of divorce (see Gottman, Murray, Swanson, Tyson, & Swanson, 2002). Hence, the goal of intervention with distressed couples is to reduce emotional inertia. Figure 3 is a summary of the means pre and post intervention for each of the five groups. For the wife, only the BOTH PLUS THERAPY intervention significantly reduced her emotional inertia from .52 to .43, t(19) = 2.43, p = .025. For the husband, only the IMPROVE FRIENDSHIP intervention reduced his emotional inertia, albeit non-significantly from .56 to .50, t(15) = 1.78, p = .096.

Insert Figure 3 about here

Uninfluenced Steady States. These parameters assess how positive or negative each person was independent of partner influence processes. This parameter is an assessment of what each person brings to the interaction, which is a function of both personality and past relationship history. These data are displayed as Figure 4. For husbands, the BOTH+THERAPY intervention resulted in a significant increase in the positivity of the uninfluenced steady state, from –1.23 to .31, t(19) = 2.57, p = .019, while the IMPROVE FRIENDSHIP intervention resulted in a nonsignificant change, t(15) = 1.06, n.s. For wives, the BOTH+THERAPY intervention resulted in a significant increase in the positivity of the uninfluenced steady state, from –1.40 to -.47, t(19) = 1.82, p = .019, while the BOTH
intervention resulted in a non-significant change, \( t(18) = .69, \) n.s.

The Attractors: Influenced Steady States. These parameters assess how positive each person was influenced by just the partner influence processes. These parameters are the “attractors” for the couple, which assess how positive each person was after influence processes occurred. To assess this change we examined whether the intervention changed the number of stable steady states in either the negative-negative quadrant of phase space, or in the positive-positive quadrant of phase space. The parameter assess the number of both husband and wife states in each quadrant (negative-negative and positive-positive). These data are displayed as Figure 5. For the negative-negative quadrant, the IMPROVE FRIENDSHIP intervention resulted in a significant reduction from 1.88 to 1.06 negative steady stable states, \( t(15) = 2.55, p = .022; \) the BOTH+THERAPY intervention also resulted in a significant reduction from 1.35 to .65, \( t(19) = 2.21, p=.040. \) For the positive-positive quadrant, only the BOTH intervention resulted in a significant increase from .11 to .42 positive steady stable states, \( t(18) = 2.36, p = .030 \) (the BOTH+THERAPY intervention and the BIBLIOTherapy intervention were not significant, \( t(19) = 1.19, \) and \( t(17) = 1.30, \) respectively).

Influence functions’ linear slopes. The influence functions describe how each person
influenced their partner across the entire range of affect. These functions are described (in part) by two numbers: the linear slope of the influence functions in the positive affect ranges and the linear slope of the influence functions in the negative affect ranges. For the wife, there were no significant changes in any of the positive slope parameters for any of the interventions. For the negative slope, two interventions produced a significant change, the IMPROVE FRIENDSHIP intervention reduced the negative slope from .67 to .35, t(15) = 3.03, p = .009, and the MANAGE CONFLICT intervention reduced the negative slope from .52 to .37, t(16) = 2.24, p = .040. Thus two of the interventions changed the basic shape of the wife’s influence functions. For the husband, the only significant effect was for the positive slope parameter, and only for the BIBIOTHERAPY intervention, which moved the positive affect slope parameter from .27 to .46, t(13) = 2.60, p = .022, thus increasing his ability to influence his wife with positive affect. These data are displayed as Figure 6.

Influence functions’ effectiveness of repair attempts. This parameter assesses how successful each person was at improving the communication once it became negative. For wives, the only intervention to significantly change the effectiveness of repair was BOTH, but unfortunately it made repair less effective, moving the amplitude of repair from 5.52 to 2.92, t(17) = -2.24, p = .039. For husbands, two interventions significantly changed the effectiveness of repair, both making repair less effective, IMPROVE FRIENDSHIP moved it from 3.93 to 2.13, t(15) = 2.80, p = .014, and BOTH, which moved it non-significantly from 3.12 to 2.36, t(18) = 2.00, p = .061, again making repair less effective. We then tested,
post-hoc, for an interaction effect of time by these two interventions. If these two interventions became a treatment factor (1= BOTH, 2= BOTH+THERAPY), a repeated measures analysis of variance found that the time-by-treatment interaction was non-significantly statistically significant, $F(1, 35) = 3.64, p =.065$. These data are displayed as Figure 7.

Influence functions’: Trigger threshold of negativity for repair onset. This parameter markes the onset of repair once the interaction has become sufficiently negative. No intervention significantly altered the threshold of negativity. Because of the pattern of results, we tested, post-hoc, for an interaction effect of time by the BOTH and BOTH+THERAPY interventions. We made these two interventions a treatment factor, which we denote “BOTH/BOTH+THERAPY” (1= BOTH, 2= BOTH+THERAPY). In that case, our repeated measures analysis of variance found that the time-by-treatment interaction was statistically significant, $F(1, 35) = 4.66, p =.038$. These data are displayed as Figure 8.

Influence functions’: Damping parameters. There were no significant intervention effects for the wife’s or husband effectiveness of damping, nor for the wife’s positive score that triggered the onset of damping. For the husband’s positive score that triggered the onset of damping, there was a non-significant significant effect for the MANAGE CONFLICT
intervention, \( t(14) = 1.90, p = .078 \), with pre-intervention score of 6.32 and post-intervention score of 4.00. Managing conflict has reduced the positivity of the onset of damping.

**Discussion**

Many of these brief interventions had a significant effect on the second of two conflict discussions. The IMPROVE FRIENDSHIP intervention accomplished the following: (1) the intervention reduced the number of influenced stable steady states in the negative-negative quadrant; this means that their interaction became far less prone to be drawn toward negative cycles; and (3) the intervention changed the negative affect portion of the wife’s influence function so that it was less steep; this means that the wife’s negative affect had a less negative subsequent result from the husband, decreasing the reciprocation of negativity. Therefore, we found that improving friendship alone had a significant impact in major ways on the nature of the second conflict discussion, particularly on reducing negativity and its reciprocation.

The MANAGE CONFLICT intervention changed the negative affect portion of the wife’s influence function so that it was less steep. This means that the wife’s negative affect had a less negative subsequent result from the husband, decreasing the reciprocation of negativity. This is a desirable result, because it is less likely to result in negative-negative chains of behavior.

Comparing the IMPROVE FRIENDSHIP intervention with the MANAGE CONFLICT intervention, we did find that the IMPROVE FRIENDSHIP intervention had a more powerful and pervasive effect on the second conflict discussion than did the MANAGE CONFLICT intervention. Improving friendship alone affected more aspects of negativity than managing conflict alone. This was an effect we predicted theoretically (see
Gottman & Silver, 1999), and we were delighted to find empirical support for our prediction.

However, what is perhaps more interesting is that the BOTH intervention accomplished something entirely new. This was our hope in designing these interventions, affecting the attractiveness of a positive affect steady state during conflict discussions. The BOTH intervention created more influenced stable steady states in the positive-positive quadrant. This means that both husband and wife became far more prone to be drawn toward positive cycles. Here is a major advantage of the systems thinking of von Bertalanffy. Not only has positive affect been increased, but a likely effect of the intervention is that a positive affect attractor has been added. Positive affect is very important in marital interaction. Recall that positive affect in our newlywed study (Gottman, Coan, Carrere, & Swanson, 1998) was the single best predictor of both marital stability and happiness. This was one of our major puzzles, namely, how is one to increase the likelihood of positive-positive affect. However, the increase in positive-positive affect steady states (attractors) means even more than just an increase in positivity. It means that positivity has become a much more likely attractor for the couple’s second conflict discussion. It has become a stable aspect of their conflict discussions. If this pattern lasts, they will therefore be drawn increasingly toward positive-positive cycles in their conflict discussions.

The BOTH+THERAPY intervention reduced the number of influenced stable steady states in the negative-negative quadrant and it also enhanced the positivity of the husband’s and wife’s influenced stable steady states. This means that again there is an increase in positivity of the second conflict discussion. But, to reiterate, it also means more than that, it means that positivity has become a much more likely attractor for the couple’s second
conflict discussion. If this pattern lasts, once again with this intervention they will therefore be drawn increasingly toward positive-positive cycles in their conflict discussions. This is a much more important goal than merely increasing positive affect during conflict. It means that the system of equations will draw them repeatedly toward a positive conversational outcome whenever they discuss a disagreement.

Even the BIBLIOTHERAPY intervention was an effective intervention for creating proximal change; it increased the positive affect slope parameter from .27 to .46, thus increasing the husband’s ability to influence his wife with positive affect. This means that she is more likely to respond positively to his positivity, increasing the likelihood that they will get into positive affect exchanges in their second conflict discussion simply by the relatively inexpensive intervention of the couple having read the Gottman & Silver (1999) book.

We were disappointed that the repair component of the influence function, both its effectiveness and threshold, were either not affected at all or negatively impacted by the interventions. The significant interaction with BOTH/BOTH+THERAPY and time suggests that merely the expectation of receiving therapy has the effect of wives tolerating more negativity from their husbands before initiating corrective repair. Perhaps wives are more confident that these issues will be dealt with in therapy, so they become less active in intervening when their husbands behavior becomes negative. This is not an effect we desired, and in subsequent interventions we may wish to address this problem directly so that repair is initiated at lower levels of negativity for wives who are expecting marital therapy. However, for the wife’s repair effectiveness, the nearly significant interaction with BOTH/BOTH+THERAPY and time suggests that just the expectation of receiving therapy
may have the effect of wives’ repair effectiveness increasing, even though they do not intervene until a higher level of husband negativity. Perhaps wives are more confident that these issues will be dealt with in therapy, so they become less active in intervening when their husbands behavior becomes negative. Given that negativity tends to increase as the conflict discussion proceeds, these wives probably wait longer to intervene, but when therapy is anticipated, their intervention attempts are more successful than when therapy is not expected. In general, these interventions appear to have a greater effect on wives directly and on the relationship, than on husbands directly. This should be an area for future investigation and intervention development.

Conclusion. The main point of this article was to suggest a new methodology for bringing the experimental study of marital interaction into the social psychology laboratory, where specific processes could be studied with very specific interventions. We proposed a measurement system for these proximal change experiments that is based on mathematical modeling of the kind that von Bertalanffy imagined when he described general systems theory. We showed that not only will some interventions decrease negative affect and negative affect reciprocity, and increase positive affect, but the BOTH intervention actually created a positive attractor for the couple, which is a worthy goal of intervention.

Placing marital interaction research on an experimental and mathematical footing by bringing the study of interactions and change into the social psychology laboratory, as well as the clinical laboratory, should accelerate our knowledge of how to effectively change interactive systems in families. We suggest that it is a methodology that could help build a library of interventions designed to change specific aspects of couples’ interaction and potentially assist in reversing or preventing destructive interaction patterns.
What is our next step in this mathematical modeling of social interaction? We are now moving from a difference equation to an ordinary differential equations (ODE) model. This innovation is possible because we now have 150 data points for every 15-minute interaction. There are many advantages of the ODE model, including estimation procedures that do not require a subset of points for which we need to assume that the influence functions are zero. We are also extending the model to include mother-father-baby triads, in which three equations are necessary to model the interacting system.
References


Footnotes

1 If the equation is $dN/dt = f(N)$, and there is a steady state solution (of $f(N) = 0$) at $N^*$, then the steady state is stable if derivative of $f$ at $N^*$ is $<0$, and unstable if it is $>0$. For a proof, see Gottman, Murray, Swanson, Tyson, & Swanson (2002), p.79.
**Figure Captions**

**Figure 1.** Bilinear form of the influence functions, with damping and repair terms

**Figure 2.** Null clines (solid line) and steady states (circles) for the model without repair or damping ($R_i=\overline{D}_i=0$).

**Figure 3.** Effects of the Interventions on Wife and Husband Inertia. (Control = BIBLIOTHERAPY)

**Figure 4.** Effects of the Interventions on Wife and Husband Uninfluenced steady state. (Control = BIBLIOTHERAPY)

**Figure 5.** Effects of the Interventions on the number of negative-negative and positive-influenced positive steady states. (Control = BIBLIOTHERAPY)

**Figure 6.** Effects of the Interventions on the Influence Functions: Wife Negative Slope and Husband Positive Slope. (Control = BIBLIOTHERAPY)

**Figure 7.** Effects of the Interventions on the Effectiveness of Repair: Husband and Wife. (Control = BIBLIOTHERAPY)

**Figure 8.** Effects of the Interventions on the negativity threshold: Husband Negative Score that Triggers Repair. (Control = BIBLIOTHERAPY)
Proximal Change Experiments with Couples

Figure 1.

Impact on Wife Positive

Husband Positive

Negative Slope

Positive Slope

Husband Negative

Impact on Wife Negative

Figure 2.
Figure 3.
Figure 4.
NEGATIVE-NEGATIVE QUADRANT

#STABLE STEADY STATES

PRE    POST

<table>
<thead>
<tr>
<th>GROUP</th>
<th>FRIENDSHIP</th>
<th>CONFLICT</th>
<th>BOTH</th>
<th>BOTH+THERAPY</th>
<th>CONTROL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated Marginal Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.0</td>
<td>1.8</td>
<td>1.6</td>
<td>1.4</td>
<td>1.2</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Figure 5.
WIFE NEGATIVE SLOPE

PRE                                      POST

Estimated Marginal Means

GROUP

FRIENDSHIP

CONFLICT

BOTH

BOTH+THERAPY

CONTROL

HUSBAND POSITIVE SLOPE

PRE                                      POST

Estimated Marginal Means

GROUP

FRIENDSHIP

CONFLICT

BOTH

BOTH+THERAPY

CONTROL

Figure 6.
HUSBAND EFFECTIVENESS
OF REPAIR

Estimated Marginal Means

GROUP

FRIENDSHIP

CONFLICT

BOTH

BOTH+THERAPY

CONTROL

PRE

POST

WIFE EFFECTIVENESS
OF REPAIR

Estimated Marginal Means

GROUP

FRIENDSHIP

CONFLICT

BOTH

BOTH+THERAPY

CONTROL

PRE

POST

Figure 7.
Figure 8.