

Randomized Controlled Trial of Imago Relationship Therapy: Exploring Statistical and Clinical Significance

Nathan C. Gehlert^a, Christopher D. Schmidt^b, Victoria Giegerich^c, and Wade Luquet^d

^aDepartment of Counseling, John Carroll University, University Heights, Ohio, USA; ^bCounseling and Education Department, Villanova University, Villanova, Pennsylvania, USA; ^cCounselor Education and Supervision Program, Kent State University, Kent, Ohio, USA; ^dBehavioral and Social Sciences Division, Gwynedd Mercy University, Gwynedd Valley, Pennsylvania, USA

ABSTRACT

For decades, couples around the world have used Imago relationship therapy (IRT) to improve their relationships. While anecdotal success stories abound, no randomized controlled trial of IRT's impact has been accomplished until now. The authors review the results of a randomized controlled trial of distressed, treatment-seeking couples who completed 12 sessions of IRT and the impact their involvement had on their marital satisfaction. Results showed that (a) individuals in the treatment condition experienced statistically significant increases in marital satisfaction, while couples in the control group did not; (b) levels of marital satisfaction did decrease significantly from posttreatment to follow-up but remained significantly higher than at pretreatment; (c) though statistically significant, the improvements experienced by the treatment group were not clinically significant improvements; and (d) while approximately one-third of participants achieved recovery during treatment, at the dyad level, only one couple achieved recovery. Further analysis and recommendations for future research are discussed.

KEYWORDS

Imago relationship therapy; couples counseling; clinical significance; randomized controlled trial

Committed romantic relationships have a significant impact on an individual's psychological and physical well-being, and it is commonly accepted that personal happiness is highly associated with the type of relationship one has with one's intimate partner (Dyrdal, Roysamb, Nes, & Vitterso, 2011; Headey, Veenhoven, & Wearing, 1991). A positive romantic relationship can buffer against mental health issues (Seikkula, Aaltonen, Kalla, Saarinen, & Tolvanen, 2013) and a significant body of scholarship has demonstrated that couples therapy is an effective method for reducing both relational distress and individual psychological symptoms (Beach, Dreifuss, Franklin, Klamen, & Gabriel, 2008; Snyder, Castellani, & Whisman, 2006; Snyder & Halford, 2012).

There are a number of approaches to couples counseling that are currently being used by therapists. Two of these approaches, behavioral couples therapy and emotion-focused couples therapy, have received the widest attention in the research literature, having been evaluated through multiple clinical trials and having been found to be effective in producing clinically and statistically significant reductions in relationship distress (Lebow, Chambers, Christensen, & Johnson, 2012; Snyder et al., 2006). Other approaches to couples therapy, such as interpersonal, cognitive-systemic, and communication-focused therapies, are currently building a research base (Snyder & Halford, 2012). However, some approaches that are experiencing widespread domestic and international use have yet to be rigorously examined.

The present randomized controlled trial seeks to bring comprehensive empirical attention to Imago relationship therapy (IRT; Hendrix, 2008), an integrative approach to couples therapy that has been practiced extensively for more than 30 years. IRT is a theoretical and applied methodology for working with couples in committed relationships. While preliminary research has been conducted on IRT, no randomized controlled trial has been completed until now. For couples and for the field of mental health, there is a critical need and great utility in evaluating the effectiveness of interventions already in broad use (Stratton, 2011). Taking into account IRT's widespread attention and use, Lazarus (2000) highlighted the need to move beyond IRT's foundation of "charisma, conjecture, anecdotes, and untestable theories" (p. 224). Additionally, Berger (1997) challenged the Imago community to conduct robust research that would include control groups, randomization, objective measures, and posttreatment follow-up.

Couple Distress

As stated previously, romantic relationships impact individual well-being both psychologically and physically (Snyder & Halford, 2012). Positive couple relationships can be a source of support to manage stress and can positively impact personal happiness (Hilpert, Bodenmann, Nussbeck, & Bradbury, 2012). The vast majority of people around the world continue to choose to marry (United Nations Social Affairs Population Division, 2003), despite falling marriage rates in recent decades and the belief of 39% of Americans that marriage is becoming obsolete (Pew Research Center, 2010). Perhaps the most relevant indicator of relational distress is the 40% to 50% divorce rate (Snyder & Halford, 2012; Stanley, 2007).

Distress in romantic relationships is itself strongly associated with mental health problems such as depression, anxiety, and substance abuse (Whisman, 2007). Such distress is also associated with poor physical health (Kiecolt-Glaser & Newton, 2001). Additionally, parents in distressed relationships are more likely to use negative parenting techniques with their children, leading to a host of potential problems, such as having a higher risk for poor mental health and lower academic achievement (Afifi, Boman, Fleisher, & Sareen, 2009; Krishnakumar & Buehler, 2000; Potter, 2010).

Couples Therapy

Couple communication styles are directly associated with marital satisfaction (Heyman, 2001; Walitzer, Dermen, Shyhalla, & Kubiak, 2013), and negative communication patterns are a strong predictor of marital dissatisfaction and marital dissolution (Kiecolt-Glaser, Bane, Glaser, & Malarkey, 2003). Therefore, most approaches to couples therapy focus on enhancing communication and cultivating the couple's emotional bond (Reibstein & Burbach, 2013).

The evidence for the efficacy of couples counseling continues to grow. In the United States, the National Registry of Evidence-based Programs and Practices, part of the United States' Substance Abuse and Mental Health Services Administration, includes several forms of couples therapy and programs for marriage enrichment (NREPP, 2015). The proliferation of evidence-based couples counseling underscores the ultimate need for therapeutic modalities to demonstrate their clinical efficacy and serves as a justification for the present study.

Imago Relationship Therapy

Harville Hendrix and Helen LaKelly Hunt developed IRT in 1980 as a theoretical and applied methodology for working with couples in committed relationships (Martin & Bielawski, 2011). Today, there are more than 1000 certified Imago therapists in more than 30 countries (Imago Relationships International [IRI], n.d.). IRT integrates psychodynamic approaches (e.g., ego psychology, attachment theory, and object-relations psychology), transactional analysis, and cognitive-behavioral approaches and hypothesizes that unconscious factors play a significant role in mate selection and the development of conflict in romantic relationships (Zielinski, 1999). Unconscious partner selection creates an opportunity to heal a connection that was lost in childhood by increasing empathy, understanding, and communication with one's adult romantic partner (Love & Shulkin, 2001). In healing childhood wounds, IRT emphasizes growth within a relational paradigm by focusing on the self-in-relation rather than the self-as-independent. Growth is seen as occurring through relationships, as opposed to through individuation and separateness, which is often touted as the pinnacle of personality development (Banks, 2011).

The self-in-relation first occurs in infancy between child and caretaker. Within this first intimate relationship, the child learns to define the self through actions and words that receive either validation or neglect from early caretakers. These interactions, in turn, facilitate a growth process that can build connection and empathy or foster defensive disconnection (Jordan, 1995). If an individual experiences ongoing violations in close relationships, then self-protection is learned and implemented (Jordan, 1995). IRT seeks to correct developmental stumbling blocks and childhood wounds by restoring the connection between partners.

IRT Interventions

Imago therapists actively help couples learn and apply connection-building skills through a number of specific interventions, such as the couples dialogue, parent-child dialogue, behavior change request dialogue, and Imago workup. The following paragraphs summarize these interventions and are drawn from the *Imago Training Manual* (IRI, 2014), part of the educational materials used by the therapists in the present study during their training.

The Couples Dialogue

Imago therapy is perhaps best recognized by the use of the couples' dialogue. The couple learns to effectively communicate by taking turns as the "sender" or the "receiver." By using a three-step process of mirroring, validating concerns, and expressing empathy, couples practice paraphrasing, interpreting content and meaning, and asking for clarification. Couples learn to express genuine care for each other and are curious about each other's views, which creates feelings of safety, even in times of disagreement.

The Parent-Child Dialogue

This dialogue takes the sender back to their experiences in childhood, allowing the sender to identify his or her thoughts and feelings associated with a childhood caretaker and then direct them towards his or her current romantic partner. The dialogue is designed to enable the receiver to experience empathy for the sender's unmet childhood needs and understand how they relate to present needs in the relationship.

The Behavior Change Request Dialogue

This process is a formal expression from the sender that allows the receiver to hear and empathize with a present frustration in the relationship and how it relates to an unmet childhood need. At the end of the dialogue, the sender requests three specific, small behavior changes that relate to the frustration (e.g., "I request that you make dinner for me once during the next week"). The receiver then chooses to try one of the requested behavior changes. The couple is taught specific goal-setting techniques to meet the expressed needs and encouraged to display gratitude for the vulnerable expression of personal needs.

The Imago Workup

The Imago workup is a psychoeducational exercise that encourages individuals to identify positive and negative traits in their partner that are similar to those of an early childhood caretaker (e.g., available, energetic, short-tempered, or overbearing). This helps the couple understand the similarities between their romantic partner and childhood caretakers and how these similarities can contribute to relationship frustrations.

Imago Research

Several nonrandomized, noncontrolled preliminary research studies have been conducted that lend some validity to the efficacy of IRT (aHannah, Luquet, & McCormick, 1997; Hannah et al., 1997b; Luquet & Hannah, 1996). IRT is usually delivered through either traditional in-office therapy or through the Getting the Love You Want Workshop (GTLYW Workshop), a manualized, 2-day psychoeducational workshop conducted by certified presenters. Studies have been conducted in both settings (Pitner & Bailey, 1998; Weigle, 2006). We present here a review of research on the use of IRT in clinical settings. For a review of preliminary studies on the GTLYW workshop, see Hogan, Hunt, Emerson, Hayes, and Ketterer (1996) and Schmidt, Luquet, and Gehlert (2015).

Luquet and Hannah (1996) hypothesized that IRT would have a positive effect on communication skills and specifically that IRT would promote empathy, intimacy, and conflict resolution in couples' relationships. The researchers administered the Marital Satisfaction Inventory (Snyder, Wills, & Keiser, 1981) to analyze couples' progress. Upon completing a manualized six-session course of IRT, the nine couples showed significant improvement on the subscales of Global Distress, Affective Communication, and Problem Solving Communication. In a further examination of data analyzed by Luquet and Hannah (1996), Hannah et al. (1997b) assessed participants' functioning in the life areas of family, health, intimacy, social life, and work. In the same sample, there were significant changes on the Well-Being, Symptoms, and Life-Functioning subscales.

Both of these studies were limited by their extremely small sample size; use of a mostly Caucasian, middle-class, and middle-aged sample; and reliance on only correlational analyses when examining associations between the outcome measures. The Hannah et al. (1997b) study was limited by its reliance on data collected using COMPASS, which has been only nominally examined in the research literature.

Hannah et al. (1997a) investigated the association between short-term IRT and outcomes of health and psychosocial wellness, narcissism, relationship maturity, didactic adjustment, and the use of Imago skills. In 21 couples, results revealed statistically significant changes from pretreatment to posttreatment on indices of dyadic adjustment, commitment, relationship maturity, and the use of Imago skills. There was also a statistically significant increase in participants' scores on the Well-Being COMPASS scale; scores on Life Functioning were not significantly higher posttreatment.

As in the aforementioned studies, the research by Hannah et al. (1997b) was limited in the use of a similarly nondiverse sample and reliance on COMPASS. Further, the study lacked *a priori* hypotheses and neglected to examine the relationships among the various outcome measures. All these studies were limited by the lack of control group data. The absence of control and randomization make it difficult, if not impossible, for researchers to determine if an intervention is responsible for change (Schulz, Chalmers, Hayes, & Altman, 1995). As stated by Hannah et al. (1997b), the data collected for these studies "can best be described as pilot data" (p. 87).

The Present Study

In response to the existence of only preliminary research on the efficacy of IRT and the relative widespread use of Imago among clinicians globally, the present randomized controlled trial sought to examine the efficacy of a 12-session course of IRT treatment with couples experiencing distress in their relationships. The independent variables were treatment condition (i.e., treatment or control group) and time in treatment. The dependent variable was relationship satisfaction. We hypothesized that:

- H1: Subsequent to treatment, there would be a main effect for time in treatment; averaging across treatment condition, the mean of the participant's relationship satisfaction scores would increase over time.
- H2: Subsequent to treatment, there would be a treatment condition by time interaction. We expected to find that over time we would find differences in mean levels of relationship satisfaction between the treatment and control groups.
- H3: Only the treatment group would experience a statistically significant increase in marital satisfaction scores over time.
- H4: At the 12-week follow-up, the treatment group would not exhibit a decrease in level of relationship satisfaction.
- H5: There would be a clinically significant increase in the level of relationship satisfaction in the treatment group.

The final hypothesis about clinically significant change was a major focus of the present study. Unlike statistical significance, clinical significance refers to the importance of determining if a change makes a real difference in the individual's life (Kazdin, 2003). For our purposes, we wanted to know not only if participants would experience statistically significant increases in relationship satisfaction, but, more importantly, if improvements would move participants from relationship discord to satisfaction (i.e., recovery). This study was the first study of IRT to examine clinically significant change.

In our analysis of clinically significant change, we would use Jacobson and Truax's (JT; 1991) method for calculating clinical significance. In the JT approach, clinically significant change is indicated when the level of functioning on an assessment post-treatment places the individual closer to the mean of the well-adjusted population than it does to the mean of the maladjusted population. Jacobson and Truax present two other measures of clinical significance but argue for the use of this least-arbitrary measure when population norms are available. Moreover, in a comparative analysis of multiple approaches to measuring clinical significance, Bauer, Lambert, and Nielsen (2004) support the JT method because it is widely-used and balances the benefits and drawbacks of other methods.

Our analysis of significant change would also include a *post hoc* analysis to determine the rate of improvement and recovery for participants in the treatment group.

We used a reliable change index (RC), which was developed by Jacobson, Follette, and Revenstorf (1984) and later amended by Christensen and Mendoza (1986):

$$RC = \frac{X_2 - X_1}{S_{diff}}$$

where X_1 represents the participant's pretreatment score, X_2 represents the post-treatment score, and S_{diff} is the standard error of difference between the two scores. A posttreatment RC larger than 1.96 is likely to occur only when the participant has experienced real change.

Method

The Sample

Participants were recruited via convenience and snowball sampling. Recruitment was managed by the researchers in a city that was not one of the cities where participants were being recruited. This allowed us to do our best to standardize recruitment in all geographic locations. Using the Internet, we identified civic organizations, graduate-level mental health programs, houses of worship, and mental health professional organizations in the geographic areas we were targeting and contacted them via e-mail requesting that our solicitation for participants be distributed to their members and contacts. We also posted solicitations on online professional and community listservs and forums. The solicitations provided a link to the study website where they could read more about the study and initiate the screening process. The primary benefit described to participants was the possibility of improving the quality of their relationship. The primary risk we described was the possibility of their relationship satisfaction decreasing and that this would be more likely to happen through participation in the control group. All participants were adults in heterosexual relationships that resided in eight metropolitan cities in the United States as well as one in Canada. The inclusion criteria consisted of the following: be cohabitating for a minimum of 1 year, have no immediate plans to terminate the relationship, have received no psychiatric treatment within the last 2 years, be free of alcohol or drug problems and primary sexual dysfunction, have no evidence of active partner abuse, not be presently involved in other psychologically oriented treatment, not be incarcerated, and be experiencing distress in their relationship. Distress was assessed using the Marital Adjustment Test (MAT; Locke & Wallace, 1959), and at least one partner was required to have a score below 100 on the assessment; there were exclusion criteria involving exceptionally low MAT scores. If couples were eligible to participate, then they were randomly placed into either the treatment or control group using an online random number generator.

Three-hundred forty individuals completed the initial online screening. Of these, 62 were individuals whose partner did not complete the screening, so they were disqualified from participating. Of the remaining 278 individuals (139 couples), 104 couples were not enrolled because one or both of them did not meet the following

inclusion criteria: 33 did not meet distress criteria; 17 qualified, but did not respond to our communications about next steps for enrollment; 17 were not available during hours when treatment would be provided in intervention condition; 11 were geographically too distant from treating therapists; 2 had previous exposure to IRT treatment or literature; and 24 for miscellaneous reasons (e.g., being on medication, currently receiving treatment, abuse, etc.).

Nineteen couples were initially enrolled in the control group. One of these couples withdrew from the study after completing the first assessment and two more did not complete the final assessment. In the treatment group, 16 couples were enrolled and two couples dropped out. One stopped attending treatment and the other did not complete assessments at the final two data collection time points. A Mann–Whitney *U* test comparing means on MAT did not reveal a significant difference between participants who did and did not complete the study. Subsequently, the sample for the present study was composed of 32 participants (16 couples) in the control group and 28 participants (14 couples) in the treatment group.

The mean age for the sample was 45 years, and ages ranged from 25 to 70 years with the majority of participants being in their 30 s (32%) or 40 s (32%). The length of the couple relationships ranged from 2 to 45 years with most of the couples having been together for 5 to 10 years (30%). Ninety-five percent of the sample completed at least some college, with 35% having completed a bachelor's degree and 37% having completed additional graduate or professional studies. Forty-seven percent reported a combined household income up to \$100,000 and 42% reported making between \$100,000 and \$200,000. Fifty percent of the couples made between \$60,000 and \$140,000. The participants were 81% white, 17% black, and 2% of Asian descent. Thirty percent of the participants had engaged in couples therapy at some point during the past 5 years and 22% had engaged in either individual or group therapy.

The pretreatment mean score on the MAT (Locke & Wallace, 1959) was 73.45 ($SD = 23.85$), indicating that the sample was indeed experiencing marital distress. The MAT, in part, assesses disagreement in several domains related to romantic relationships on a 6-point Likert-type scale ranging from always agree to always disagree. We conceptualize a high level of disagreement as either frequently disagree, almost always disagree, and always disagree. High levels of disagreement were experienced in our sample in the following domains: 52% sex relations, 44% demonstration of affection, 40% handling finances, 36% matters of recreation, 26% conventionality (right good, proper conduct), 23% dealing with in-laws, and 20% friends. Based on a frequency analysis of presenting problems covered in the screening assessment, there were no observable differences in presenting problems between the intervention and control groups. Further, 55% of participants indicated that they were less than happy with their relationship and 60% reported that they at least occasionally wished they were not in the relationship. There were no significant differences between the treatment and the control groups with respect to MAT scores. There

were also no significant group differences on characteristics such as age, length of relationship, or level of education.

Procedures and Measures

Data for this study were collected online using PsychData. Initial enrollment screening was conducted online to determine if prospective participants met the eligibility requirements. Participants completed an informed consent and the screening questions separately from their partner at a time of their choosing. The informed consent included information on the funders of the study. The participants were not paid for participating in the study.

Participating couples in the treatment group committed to completing 12 treatment sessions, 90 minutes in length each, at no cost. The 12 sessions had to occur within a timeframe of 18 weeks and the interval between sessions could not exceed 2 weeks. They also committed to not engaging in treatment during the follow-up period. Given the lack of clinical evidence supporting IRT, the length of treatment for the study was somewhat arbitrary. A review of couples therapy research literature indicated that 12 sessions was the mode number of sessions for couples (Fals-Stewart, Birchler, & O'Farrell, 1996; Greenberg, Warwar, & Malcolm, 2010; Schade et al., 2014; Tilden, Gude, Sexton, Finset, & Hoffart, 2009; Trudel et al., 2008). This was a number also frequently suggested by IRT therapists with whom we consulted and fit within our research budget.

We assessed couples pretreatment, mid treatment, and posttreatment, which was also common practice in the literature we reviewed. We did not find a clear best-practice for length of assessment follow-up posttreatment, so we selected a 12-week follow-up to roughly mirror the length of treatment. The treatment group was assessed before the first counseling session (T_1), after the sixth session (T_2), after the twelfth (i.e., final) session (T_3), and 12 weeks after the final session (T_4). Participants in the control group were assessed on the same schedule as those in the treatment group: at study enrollment (T_1), 6 weeks later (T_2), and again 12 weeks after enrollment (T_3). Control group participants in all cities were offered free admission to a GTLYW Workshop after they completed the 12-week assessment. The workshop was intended to serve as an incentive for participating in the control group. Because of the option to participate in the workshop immediately after treatment, control group participants were not assessed at a 12-week follow-up. We were unable to manipulate the offering of the workshop to fall after a 12-week period. Over the course of 12 weeks, the control group participants read *You Just Don't Understand: Women and Men in Conversation* (Tannen, 2007). They received no other interventions, and compliance with the reading exercise was not assessed.

The clinicians administering treatment were certified Imago relationship therapists who were also either Imago faculty or Imago consultants (i.e., the highest level of training possible as an Imago clinician). They were recruited by e-mailing Imago faculty and consultants and requesting their participation. Some were recruited by in-person or phone requests. The therapists were all licensed in their respective

fields (e.g., counseling, social work, and psychology) within the state they currently practice. They volunteered to participate in the study after being solicited to participate by IRI. The therapists were all offered payment of \$100 per session. Several therapists volunteered their time and opted not to receive payment. There were one or two therapists in each of the eight cities where we collected data, and each therapist agreed to treat one or two couples. None of the therapists were authors of this study.

To ensure treatment fidelity, all sessions were video-recorded and evaluated using reactive observation (Bernard, 2012). Participants were made aware of and agreed to having their session recorded when they completed informed consent. No research tool existed for assessing fidelity of IRT treatment. Thus, the variables and behavioral indicators under observation were based off the Scoring System for Imago Therapy Certification (SSITC; IRI, 2014), which is the standardized evaluation tool for therapists becoming certified in IRT. The SSITC was adjusted for research purposes (several written-feedback items were removed) and renamed SSITC-R. The SSITC-R includes 14 Likert-type items that assess the use of IRT techniques, ability of the therapist to move the couple into deep behavior- and feeling-targeted conversation, and inclusion of psychoeducational content related to IRT principals. The rating scale was as follows: 1 = present in the video; 2 = not present in the video but not necessary given the content of the segment; 3 = not present in the video but necessary. The SSITC-R consists of three separate sections: Section A includes establishing and maintaining the structure of the IRT dialogue; Section B includes the therapist's facilitation of "deepening" techniques; Section C evaluates the inclusion of Imago psychoeducation. Total scores on the SSITC-R range from 14 to 42. A perfect score of 14 signifies the use of all major Imago principles and techniques, while a score of 42 indicates a complete lack of Imago principles and techniques. Scores lower than 23 indicate fidelity with IRT best practices.

The four raters who participated in the video reviews were recruited using the same procedure that was used to recruit the clinicians for the study. They were all certified IRT faculty or consultants who were not participating in the study as clinicians; they volunteered their time and were not compensated. Before beginning the rating process for the research study, all raters participated in practice reviews using the SSITC-R and video segments of couples receiving IRT (these videos were of couples outside the research study). After the first practice review was completed, the raters and researchers discussed areas of disagreement or confusion in utilizing the SSITC-R and it was adjusted as needed. The final version of the SSITC-R was established and three additional videos were rated to ensure reliability before beginning the rating process with the videos from the study participants.

Twenty-minute-long video clips were selected from the midpoint of each therapist's fifth and ninth session with a couple and rated by the reviewers using the SSITC-R. Segments were reviewed by more than one rater. The mean of the reviewed videos was 20.8, indicating that therapists in the study were delivering treatment at or above the level that is required for certification as an IRT clinician.

Table 1. Mean MAT scores across time.

	T ₁		T ₂		T ₃		T ₄ ^a	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Control group								
Raw scores	75.22	24.94	75.08	25.36	74.19	23.95		
Adjusted scores ^b	0.00	1.00	-0.01	1.02	-0.04	0.96		
Treatment group								
Raw scores	71.43	22.83	79.27	28.09	91.32	24.42	88.14	20.60
Adjusted scores	0.00	1.00	0.34	1.23	0.87	1.07	0.73	0.90

Note. MAT is Marital Adjustment Test (Locke & Wallace, 1959). T₁ is study enrollment, T₂ is week 6, T₃ is week 12 (posttreatment for treatment group), and T₄ is 12-week follow-up.

^aOnly data for the 20 participants (10 couples) who completed the T₄ are presented here.

^bThe raw MAT scores at each time point were converted to z-scores using the T₁ mean and standard deviation for each group.

Marital Adjustment Test

The MAT (Locke & Wallace, 1959) is a 15-item self-report questionnaire that is one of the most frequently used measures of relationship satisfaction. Kazak, Jarmas, and Snitzer (1988) described it as the “grandparent” of marital satisfaction measures. The scale focuses on issues such as involvement in joint activities, demonstration of affection, frequency of marital complaints, level of loneliness and well-being, and partner agreement on significant issues. Scores on the scale are calculated by summing numerical weights that correspond to each item. Higher MAT scores indicate higher levels of overall marital satisfaction. Chronbach’s α in data from the present study was .72 for T₁, .79 for T₂, .77 for T₃, and .67 for T₄. The test-retest reliability, as measured in data from the control group between T₁ and T₃, was .84.

Results

Results were calculated using IBM SPSS Statistics Version 20. Table 1 presents the mean MAT scores in the treatment and control groups longitudinally across the study. The control group exhibited slightly decreasing MAT scores, while the treatment group exhibited rising scores to T₃ and then a decline in the final MAT score at T₄. Despite the random assignment to treatment condition, there were differences between the T₁ MAT scores (control group $M = 75.22$, treatment group $M = 71.43$). To account for this variability in the dependent variable and potential variability across time, we employed Robert’s z-score method of adjustment, which is described in Hamilton et al. (1954). Separately for each group (i.e., control and treatment), the raw MAT scores at each time point were converted to z-scores using the T₁ mean and standard deviation for that group. In other words, the z-scores were scaled to T₁ scores. The adjusted scores are also presented in Table 1. Figure 1 illustrates the change in the adjusted MAT scores in both groups’ scores across time. The MAT scores for the control group declined slightly, while the MAT scores for the treatment group increased appreciably.

Treatment Effects

Using the adjusted z-scores, we conducted a repeated-measures MANOVA to test for main effects and interactions for time and treatment condition (i.e., treatment

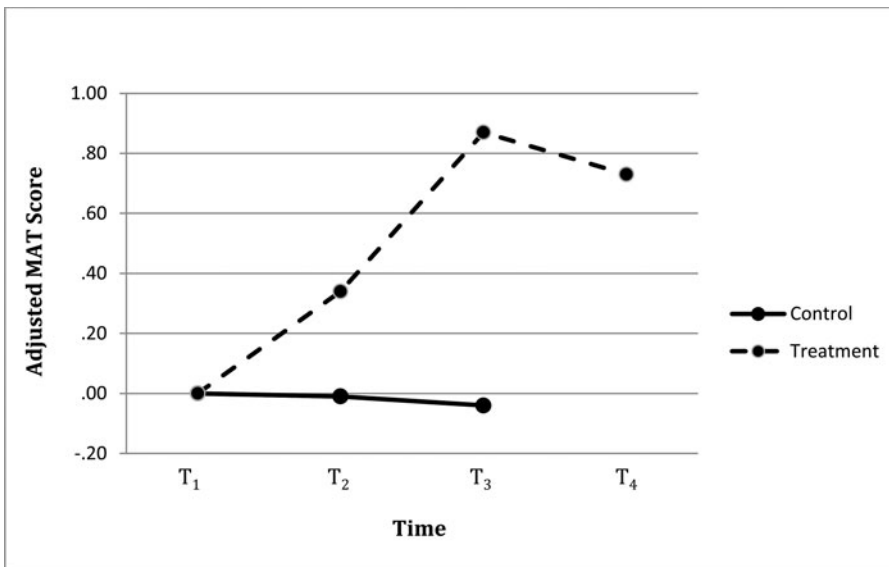


Figure 1. Mean adjusted scores on the Marital Adjustment Test (MAT; Locke & Wallace, 1959) across time.

or control). This analysis yielded a significant multivariate main effect for time, $F(2,116) = 11.91, p < .001, d = .45$. A significant main effect for time means that when ignoring the effect of treatment condition, the average of all participants' scores on the MAT increased over time. The power to detect the effect, as calculated using G*Power, was 1.00. There was no significant main effect for treatment condition, $F(1, 58) = 2.79, p = .100$, meaning that when we ignore the effect of time, treatment condition alone did produce differences in participant's level of marital satisfaction. Further, we also found a significant interaction between time and treatment condition, $F(2, 116) = 14.48, p < .001$. The significant interaction means that, in describing the main effect for time (i.e., that time has an effect on MAT), the main effect must be qualified by stating that the effect depends on treatment condition. In other words, the effect of time on MAT depends on condition. We can make the causal inference that the treatment group's MAT levels improved because they were in the treatment condition. Subsequently, hypotheses H1 and H2 were confirmed.

Increases in MAT Scores

Next, we conducted paired sample *t*-tests to determine whether the changes in MAT scores across time were significant. As can be seen in Table 2, the mean MAT score in the control group did not change significantly between any time points. Conversely, the mean MAT scores in the treatment group did significantly increase over the course of treatment from T₁ to T₃. Furthermore, in the treatment group, there was not significant change between T₁ to T₂, but there was significant improvement from T₂ to T₃. So, only the treatment group exhibited a statistically significant increase in mean MAT scores during the treatment period. H3 was confirmed.

Table 2. Paired sample *t*-tests of mean MAT scores across time.

Treatment Condition	Mean Comparison	<i>t</i>	<i>df</i>	<i>p</i>	Cohen's <i>d</i>
Control	T ₁ and T ₂	0.06	31	.956	.01
	T ₂ and T ₃	0.84	31	.406	.03
	T ₁ and T ₃	0.42	31	.681	.04
Treatment	T ₁ and T ₂	-1.81	27	.810	-.30
	T ₂ and T ₃	-5.35	27	.001*	-.46
	T ₁ and T ₃	-5.23	27	.001*	-.84
	T ₃ and T ₄ ^a	2.27	19	.035**	.31
	T ₁ and T ₄	-4.15	19	.001*	-.59

Note. MAT is Marital Adjustment Test (Locke & Wallace, 1959). T₁ is study enrollment, T₂ is week 6, T₃ is week 12 (posttreatment for treatment group), and T₄ is 12-week follow-up.

^aOnly data for the 20 participants (10 couples) who completed the T₄ assessment were used in the *t*-tests involving T₄ data.

p* < .001, *p* < .05.

Our final paired sample *t*-test checked for change in mean MAT scores for the 20 participants who completed the assessment at the 12-week follow-up. As can be seen in Table 2, there was a significant decrease from T₃ to T₄. The treatment group did exhibit a statistically significant change in mean MAT scores posttreatment. Hypothesis H₄ was rejected. However, the T₄ mean score for these individuals remained significantly higher (*SD* = .90) than at T₁.

To shed light on the practical significance and magnitude of change between time points, Table 2 also presents effect sizes. Regarding the statistically significant differences in the treatment group, the differences range in magnitude from medium (T₂ to T₃; *d* = -.46) to large (T₁ to T₃; *d* = -.84) (Cohen, 1988).

Clinically Significant Change

Figure 2 illustrates the JT approach to calculating clinically significant change with the data in the present study. In Locke and Wallace's (1959) normative data, the mean MAT score for well-adjusted couples was 135.90. In the present study, the mean MAT score at T₁ for all participants in our sample was 73.45. These are the two means against which a participant's posttreatment MAT score was compared. The midpoint between these scores is 104.68, a number that serves as the cutoff point past which a participant's MAT score must improve in order to indicate clinically significant change. In other words, a MAT score below 104.68 would suggest an individual to be a member of the population of individuals maladjusted in their romantic relationship. A score above 104.68 would likely place that individual in the population of well-adjusted individuals. As Figure 2 illustrates, the treatment group's T₃ mean MAT score of 91.32 is below this cutoff point, indicating that the treatment group did not experience a clinically significant improvement in marital adjustment. Subsequently, H₅ was not supported.

Rates of Improvement and Recovery

Our final analysis was the analysis to determine the rate of improvement and recovery for participants in the treatment group. Table 3 shows the pretreatment and

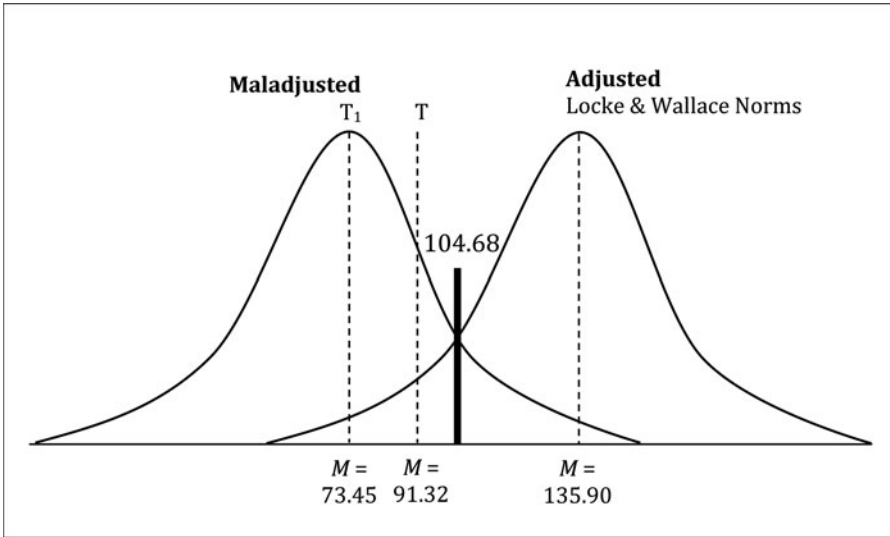


Figure 2. Mean scores on the Marital Adjustment Test (MAT) at pretreatment (T_1) compared with mean scores of adjusted individuals from Locke & Wallace’s (1959) normative data; 104.68 is the cutoff point between the two means. T_3 is the mean score of the treatment group subsequent to treatment.

Table 3. Individual MAT scores, improvement, and recovery.

Participant	Couple	T_1 MAT	T_3 MAT ^a	RC ^b	Improved but not recovered	Recovered
1	A	103	115	0.93	N	Y
2	A	94	102	0.62	N	N
3	B	95	109	1.08	N	Y
4	B	50	70	1.55	N	N
5	C	70	70	0.00	N	N
6	C	53	76	1.78	N	N
7	D	72	130	4.49	N	Y
8	D	78	105	2.09	N	Y
9	E	37	87	3.88	Y	N
10	E	72	93	1.63	N	N
11	F	65	84	1.46	N	N
12	F	93	106	1.01	N	Y
13	G	78	99	1.63	N	N
14	G	106	116	0.77	N	Y
15	H	44	32	-0.93	N	N
16	H	31	35	0.31	N	N
17	I	60	92	2.48	Y	N
18	I	44	72	2.21	Y	N
19	J	90	103	1.01	N	N
20	J	77	113	2.79	N	Y
21	K	36	121	6.58	N	Y
22	K	68	75	0.53	N	N
23	L	75	103	2.17	Y	N
24	L	110	120	0.77	N	Y
25	M	70	84	1.08	N	N
26	M	105	91	-1.08	N	N
27	N	46	54	0.58	N	N
28	N	78	100	1.70	N	N

Note. MAT is Marital Adjustment Test (Locke & Wallace, 1959). T_1 is study enrollment and T_3 is posttreatment at week 12. RC, Reliable Change Index; Y, yes; N, no.

^aBold text represents recovery. ^bBold text indicates improvement ($RC > 1.96$).

posttreatment scores for each participant, as well as their corresponding RC. Participants are also classified as improved ($RC > 1.96$) or recovered (on the basis of the 104.68 cutoff point from the clinical significance analysis). Fourteen percent of the sample experienced some improvement, while 32% experienced improvements at the level of recovery. Table 3 also shows the couples and the partners in each couple are represented by the same letter in the Couple column. It is also important to consider recovery at the dyad level, which we define as both partners experiencing recovery. By this definition, only one couple, couple D, recovered, while couples A, G, J, and L were all close to recovery.

Discussion

The main purpose of this study was to assess marital satisfaction outcomes following a 12-week course of IRT. The results showed that (a) individuals in the treatment condition experienced statistically significant increases in marital satisfaction, while couples in the control group did not; (b) levels of marital satisfaction did decrease significantly posttreatment but remained significantly higher than at pretreatment; (c) though statistically significant, the improvements experienced by the treatment group were not clinically significant improvements; and (d) while approximately one-third of participants achieved recovery during treatment, at the dyad level, only one couple achieved recovery.

The need for more robust empirical examination of IRT has been great, especially given the widespread global use of IRT and the nearly-two-decade period since the last study of IRT in a clinical setting. Our findings, which are the first to be based on treatment and control data, lend additional support to the possible efficacy of IRT as a treatment modality that benefits couples experiencing distress in their relationships. Importantly, because randomized controlled trials give researchers the ability to make casual inferences, they can provide the strongest evidence of a treatment's efficacy. This evidence is important because it is imperative for mental health practitioners to use forms of treatment with evidence to suggest that they provide the greatest chance of clinical improvement (Hunsley, Dobson, Johnston, & Mikail, 1999). While we did exclude participants with some severe presenting problems, research by Miller, Yorgason, Sandberg, and White (2003) and Whisman, Dixon, and Johnson (1997) indicates that our criteria did not exclude the majority of presenting issues that couples bring to treatment. Our research design was also strengthened by our use of multiple therapists in multiple cities across North America. Our findings represent an initial step in forming a base of evidence on which IRT clinicians can make decisions about the care of clients. As community providers and funders demand more accountability from clinicians, they will require even more evidence that builds on our findings.

Importantly, for this sample of distressed couples, significant change was not detectible at the midpoint of treatment and was only expressed after 12 sessions of treatment. This finding is inconsistent with the findings of Luquet and

Hannah (1996) and Hannah et al. (1997a, 1997b). Those authors found that a 6-week course of treatment in IRT was associated with improvements in marital functioning. An explanation for these differing outcomes may be the result of the assessments used in the various studies. The Narcissistic Personality Inventory, COMPASS, and Relationship Maturity Index are measures of various aspects of individual and relational functioning, rather than measures of relationship satisfaction. Further, COMPASS was designed to measure progress in individual therapy (Lueger, 2012). Luquet and Hannah (1996) did use the robust Marital Satisfaction Inventory but assessed only functioning on three of 11 subscales, making it difficult to grasp a true picture of relationship well-being. The only previous research of IRT to use a robust, well-validated assessment was the work of Hannah et al. (1997b), who used the Dyadic Adjustment Scale. However, the pretreatment mean score on the DAS ($M = 103.00$, $SD = 22$) in their sample was so close to the mean of well-adjusted individuals ($\bar{M} = 114.80$, $SD = 17.80$; Spanier, 1976) that they were likely sampling from the population of well-adjusted adults. Using a robust measure of relationship satisfaction and sampling from the population of distressed couples, we found that 12 sessions of treatment were necessary to produce gains in dyadic adjustment.

While statistical indications of a treatment outcome are important, they have little to do with the size, quality, or clinical significance of change. Statistical analyses shed little light on the actual efficacy of treatment. The testing of clinical significance is a critical advancement in the evaluation of interventions (Kazdin, 2003). According to Kazdin,

clinical significance refers to the practical or applied value or importance of the effect of an intervention, that is, whether the intervention makes a real (e.g., genuine, palpable, practical, noticeable) difference in everyday life to the clients or to others with whom the client interacts. (2003, p. 691)

Calculating effect sizes, such as Cohen's d , sheds some light on the magnitude of treatment effects but is limited in that the result is still a statistic that is not based on standards of efficacy that are set by consumers, clinicians, and researchers (Jacobson & Truax, 1991).

What does this mean in relation to our results? The treatment group's improvements in relationship satisfaction were detectable through our statistical analyses. The magnitude of change between T_1 and T_3 , as denoted by Cohen's d , could even be classified as large. But much like the shrinking of a tumor during chemotherapy, change, even great change, does not necessarily equate to recovery. To determine the meaningfulness of treatment, we employed a measure of clinical significance based on norms of well-adjusted individuals and found that the treatment group did not experience gains that would number them among the well-adjusted population. In other words, their improvements in relationship satisfaction would likely not have practical effects on their everyday lives. Importantly, these improvements did deteriorate significantly during the 12-week follow-up. This is unsurprising, given that recovery was not realized for so many.

Our findings suggest that 12 sessions are insufficient to produce meaningful change for most couples, though it was enough for one-third of the participants. However, at the dyad level, it was sufficient for only one couple. The fact that four other couples, 29% of the treatment group, were so close to recovery enables us to speculate that just a few more sessions were needed for a large portion of the sample. But without additional data, we can only conclude that the IRT intervention did not produce meaningful change for most couples.

Our findings also illustrate the limitation of examining statistical significance and effect sizes. In considering only those statistics, our results look quite positive. However, examining clinical significance and practical value of change provided a more nuanced and precise interpretation of these data. We acknowledge, too, that when considering clinical significance, the results are also less optimistic.

Limitations and Recommendations for Future Research

There are several important limitations of this study. First, while our sample size was typical of research on couples counseling, by statistical standards, it was still small. Because we did not employ an intent-to-treat analysis, the few dropouts in our small sample could have increased the chance of Type 1 error.

Second, in IRT, couples are often encouraged to attend the GTLYW Workshop as an adjunct to treatment. Hypothetically, this provides an initial, large dose of treatment and aids recovery. We did not include the workshop in our protocol primarily because of budget constraints. Several of the clinicians in our study reported that the participants they were treating would have benefitted greatly from attending the workshop. Future randomized controlled trials of IRT could include a third treatment condition where participants would experience in-office treatment and the workshop. Given the prevalence of referring couples to the workshop, the relationship of the workshop to therapy should be investigated.

Given the upward slope of the MAT scores for participants in the treatment group, the number of couples who were close to recovery, and the declines during the follow-up period, we could hypothesize that the treatment was effective but the dosage was not sufficient for recovery. Therefore, the length of treatment in the present study was possibly a third limiting factor. Given that recovery is the goal of treatment, future research could examine longer courses of treatment. This would shed light on this potential limiting factor and the competing possibility that IRT is an ineffective treatment.

Fourth, our findings shed light on the possible efficacy of a course of IRT treatment, but do not illuminate the efficacy of specific interventions within therapy. Future research could investigate how specific interventions contribute to outcomes.

Fifth, our inclusion criteria would be expected to influence our findings. We had restrictive inclusion criteria for participants in this study. Certainly many couples present for treatment with plans to terminate the relationship or with co-occurring psychiatric, substance abuse, sexual dysfunction issues. Further, our

sample was homogenous demographically. All participants were heterosexual and most were Caucasian, college educated, and high earners. Our use of online recruitment and computer screening required computer access and literacy; these requirements may have excluded participants with lower socioeconomic status. Our findings should be generalized only tentatively to other populations; further research with diverse populations is warranted. Our inclusion criteria permitted the participation of individuals who had received counseling in the past 5 years, a potential confounding variable. Thus, the present treatment could have served as a refresher for some participants, rather than an independent clinical experience. Additionally, some of our participant's T₁ MAT scores were near our exclusion cutoff point of 100. It is debatable whether or no these couples were truly distressed.

Sixth, as in many studies of marital satisfaction, the present study relied solely on self-report data. Our data are potentially biased in that there was no check on the accuracy of the self-report assessments. Ideally, participants would be interviewed or observed by a rater who was blind to treatment condition.

Seventh, despite its age, we relied on a statistic for well-adjusted couples that Locke and Wallace (1959) acquired by identifying participants “judged to be exceptionally well-adjusted in marriage by friends who knew them well” (p. 254). Marriage has certainly changed in the last half decade. Other, more current, researchers have attained samples of well-adjusted couples by sampling from the general population with the assumption that non-treatment-seeking couples are well adjusted. These samples have mean levels of marital adjustment that are lower than Locke and Wallace's (1959) statistic. Given that we focused our investigation on clinically significant change, we opted for a more rigorous cutoff point that would indicate a greater likelihood that participants passing it in treatment would indeed be recovered (i.e., well adjusted) and not simply out of the clinical population.

Finally, the community of IRT therapists and educators should put every effort into recruiting and supporting scholars among its ranks. The dearth of research and contemporaneous focus on intervention and program development over the past decades contradict the movement toward evidence-based practice. Ideally, theoretical scholarship and program development should stem from systematic research, allowing clinicians to integrate their clinical expertise with the best external evidence (Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996).

Despite the limitations discussed here, this study contributes to the field of couples therapy in important ways. There is consensus that couples therapy has positive impacts on psychological and the physical health. Nevertheless, no approach has been shown to be more effective than others (Snyder & Halford, 2012; Snyder et al., 2006). Therefore, it is vital to examine all approaches in broad use. While widespread in practice, IRT has endured despite lack of empirical validation. Our findings underscore the possible validity of IRT treatment and provide a basis for researchers to continue to evaluate the efficacy of IRT.

Funding

The study was made possible with funding from Imago Relationships International and The Sister Fund.

References

- Affi, T. O., Boman, J., Fleisher, W., & Sareen, J. (2009). The relationship between child abuse, parental divorce, and lifetime mental disorders and suicidality in a nationally representative adult sample. *Child Abuse & Neglect*, *33*, 139–147.
- Banks, A. (2011). The mythic reality of the autonomous individual: Developing the capacity to connect. *Zygon*, *46*, 168–182.
- Bauer, S., Lambert, M. J., & Nielsen, S. L. (2004). Clinical significance methods: A comparison of statistical techniques. *Journal of Personality Assessment*, *82*(2), 60–70.
- Beach, S. R. H., Dreifuss, J. A., Franklin, K. J., Klamen, C., & Gabriel, B. (2008). Couples therapy and the treatment of depression. In A. S. Gurman (Ed.), *Clinical handbook of couples therapy* (4th ed., pp. 545–566). New York, NY: Guilford Press.
- Berger, R. (1997). What can clinicians learn from research? *The Journal of Imago Relationship Therapy*, *2*(1), 71–80.
- Bernard, H. R. (2012). *Social research methods: Qualitative and quantitative approaches*. New York, NY: Sage.
- Christensen, L., & Mendoza, J. L. (1986). A method of assessing change in a single subject: An alteration of the RC index. *Behavior Therapy*, *17*, 305–308.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Dyrdal, G. M., Roysamb, E., Nes, R. B., & Vitterso, J. (2011). Can a happy relationship predict a happy life? A population-based study of maternal well-being during the life transition of pregnancy, infancy, and toddlerhood. *Journal of Happiness Studies*, *12*, 947–962. doi:10.1007/s10902-010-9238-2
- Fals-Stewart, W., Birchler, G. R., & O'farrell, T. J. (1996). Behavioral couples therapy for male substance-abusing patients: Effects on relationship adjustment and drug-using behavior. *Journal of Consulting and Clinical Psychology*, *64*(5), 959–972. doi:10.1037//0022-006x.64.5.959
- Fradenburg, L. A., Harrison, R. I., & Baer, D. M. (1995). The effect of some environmental factors on interobserver agreement. *Research in Developmental Disabilities*, *16*(6), 425–437.
- Greenberg, L., Warwar, S., & Malcolm, W. (2010). Emotion-focused couples therapy and the Facilitation of forgiveness. *Journal of Marital and Family Therapy*, *36*(1), 28–42. doi:10.1111/j.1752-0606.2009.00185.x
- Hamilton, M., Pickering, G. W., Roberts, J. A. F., & Sowry, G. S. C. (1954). The aetiology of essential hypertension : The arterial pressure in the general population. *Clinical Science*, *13*(1), 11–35.
- Hannah, M. T., Luquet, W., & McCormick, J. (1997a). COMPASS as a measure of the efficacy of couples' therapy. *American Journal of Family Therapy*, *25*, 76–91.
- Hannah, M. T., Luquet, W., McCormick, J., Galvin, K., Ketterer, K., May, K., ... Kot, L. A. (1997b). Short-term Imago therapy and changed in personal distress. *The Journal of Imago Relationship Therapy*, *2*(2), 55–65.
- Haynes, S. N., Follingstad, D. R., & Sullivan, J. C. (1979). Assessment of marital satisfaction and interaction. *Journal of Consulting and Clinical Psychology*, *47*, 789–791. doi:10.1037//0022-006X.47.4.789
- Headey, B., Veenhoven, R., & Wearing, A. (1991). Top-down versus bottom-up theories of subjective well-being. *Social Indicator Research*, *24*, 81–100. doi:10.1007/BF00292652

- Hendrix, H. (2008). *Getting the love you want: A guide for couples*. New York, NY: Owl Books.
- Heyman, R. E. (2001). Observation of couple conflicts: Clinical assessment applications, stubborn truths, and shaky foundations. *Psychological Assessment, 13*(1), 5.
- Hilpert, P., Bodenmann, G., Nussbeck, F., & Bradbury, T. N. (2012). Predicting relationship satisfaction: A matter of conflict, positivity, or support? *Family Science, 4*, 110–120. doi:10.1080/19424620.2013.830633
- Hogan, T., Hunt, R., Emmerson, D., Hays, R., & Ketterer, K. (1996). An evaluation of satisfaction with “Getting the Love You Want” weekend workshop. *Journal of Imago Relationship Therapy, 1*, 67–75.
- Hunsley, J., Dobson, K. S., Johnston, C., & Mikhail, S. F. (1999). The science and practice of empirically supported treatments. *Canadian Psychology/Psychologie Canadienne, 40*(4), 316–319. doi:10.1037/h0086849
- Imago Relationships International. (2014). *Clinical training in Imago relationship therapy: Training manual text*. Lexington, KY: Author.
- Imago Relationships International. (n.d.). Imago around the world. Retrieved from <http://pro.imagorelationships.org/ABOUT/ImagoWorldwide.aspx>
- Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy, 15*, 336–352.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12–19.
- Jordan, J. V. (1995). *Relational awareness: Transforming disconnections*. Work in progress, No. 76. Stone Center for Developmental Services and Studies, Wellesley, MA.
- Kazak, A. E., Jarmas, A., & Snitzer, L. (1988). The assessment of marital satisfaction: An evaluation of the Dyadic Adjustment Scale. *Journal of Family Psychology, 2*(1), 82–91. doi:10.1037/h0080475
- Kazdin, A. E. (2003). Clinical significance: Measuring whether interventions make a difference. In A. E. Kazdin (Ed.), *Methodological issues and strategies in clinical research* (pp. 691–710). Washington, DC: American Psychological Association.
- Kiecolt-Glaser, J. K., Bane, C., Glaser, R., & Malarkey, W. B. (2003). Love, marriage, and divorce: Newlyweds’ stress hormones foreshadow relationship changes. *Journal of Consulting and Clinical Psychology, 71*, 176–188.
- Kiecolt-Glaser, J. K., & Newton, T. L. (2001). Marriage and health: His and hers. *Psychological Bulletin, 127*, 472–503.
- Krishnakumar, A., & Buehler, C. (2000). Interparental conflict and parenting behaviors: A meta-analytic review. *Family Relations, 49*, 25–44.
- Lazarus, A. (2000). Working effectively and efficiently with couples. *The Family Journal, 8*, 222–228.
- Lebow, J. L., Chambers, A. L., Christensen, A., & Johnson, S. M. (2012). Research on the treatment of couple distress. *Journal of Marital and Family Therapy, 38*, 145–168.
- Locke, H. J., & Wallace, K. M. (1959). Short marital adjustment and prediction tests: Their reliability and validity. *Marriage and Family Living, 21*, 251–255. doi:10.2307/348022
- Losada, J. L., & Manolov, R. (2015). The process of basic training, applied training, maintaining the performance of an observer. *Quality & Quantity, 49*(1), 339–347.
- Love, P., & Shulkin, S. (2001). Imago theory and the psychology of attraction. *The Family Journal, 9*, 246–249. doi:10.1177/1066480701093002
- Lueger, R. J. (2012). The Integra/COMPASS tracking assessment system. *Integra Science and Practice, 2*, 20–30.
- Luquet, W., & Hannah, M. T. (1996). The efficacy of brief Imago therapy: Preliminary findings. *Journal of Imago Relationship Therapy, 1*, 67–75.

- Martin, T. L., & Bielawski, D. M. (2011). What is the African American's experience following Imago education? *Journal of Humanistic Psychology, 51*, 216–228. doi:10.1177/0022167809352379
- Miller, R. B. M., Yorgason, J. B., Sandberg, J. G., & White, M. B. (2003). Problems that couples bring to therapy: A view across the family life cycle. *The American Journal of Family Therapy, 31*, 395–407.
- National Registry of Evidence-based Programs and Practices (NREPP). (2015). Retrieved from <http://www.nrepp.samhsa.gov/SearchResultsNew.aspx?s=b&q=marriage>
- Pershing, J. (Ed.). (2006). *Handbook of human performance technology: Principles, practices, and potential* (3rd ed.). San Francisco, CA: Pfeiffer.
- Pew Research Center. (2010). *The decline of marriage and rise of new families*. Retrieved from <http://pewresearch.org/pubs/1802/decline-marriage-rise-new-families>
- Pitner, G. D., & Bailey, W. R. (1998). The getting the love you want workshop, marital satisfaction and relationship quality. *The Journal of Imago Relationship Therapy, 3*(1), 35–48.
- Potter, D. (2010). Psychosocial well-being and the relationship between divorce and children's academic achievement. *Journal of Marriage and Family, 72*, 933–946.
- Reibstein, J., & Burbach, F. (2013). An increasingly convincing case for couples therapy. *Journal of Family Therapy, 35*, 225–228. doi:10.1111/1467-6427.12019
- Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence-based medicine: What it is and what it isn't; it's about integrating individual clinical expertise and the best external evidence. *British Medical Journal, 312*, 71–72.
- Schade, L. C., Sandberg, J. G., Bradford, A., Harper, J. M., Holt-Lunstad, J., & Miller, R. B. (2014). A longitudinal view of the association between therapist warmth and couples' in-session process: An observational pilot study of emotionally focused couples therapy. *Journal of Marital and Family Therapy, 41*(3), 292–307. doi:10.1111/jmft.12076
- Schmidt, C. D., Luquet, W., & Gehlert, N. C. (2015). Evaluating the impact of the imago couples workshop on relational adjustment and communication patterns. *Journal of Couple and Relationship Therapy, 14*(4), 201–214. doi:10.1080/15332691.2014.978061
- Schulz, K. F., Chalmers, I., Hayes, R. J., & Altman, D. G. (1995). Empirical evidence of bias: Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *Journal of the American Medical Association, 273*, 408–412.
- Seikkula, J., Aaltonen, J., Kalla, O., Saarinen, P., & Tolvanen, A. (2013). Couples therapy for depression in a naturalistic setting in Finland: A 2-source randomized trial. *Journal of Family Therapy, 35*, 281–302. doi:10.1111/j.1467-6427.2012.00592.x
- Snyder, D. K., & Halford, W. K. (2012). Evidence-based couples therapy. *Journal of Family Therapy, 34*, 229–249. doi:10.1111/j.1467-6427.2012.00599.x
- Snyder, D. K., Castellani, A. M., & Whisman, M. A. (2006). Current status and future directions in couples therapy. *Annual Review of Psychology, 57*, 317–344.
- Snyder, D. K., Wills, R. M., & Keiser, T. W. (1981). Empirical validation of the Marital Satisfaction Inventory: An actuarial approach. *Journal of Consulting and Clinical Psychology, 49*(2), 262.
- Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *Journal of Marriage and Family, 38*, 15–28.
- Stanley, S. M. (2007). What really is the divorce rate? Retrieved from https://www.prepinc.com/docs/content/articles/what_is_divorce_rate_8-3-2007.pdf
- Stratton, P. (2011). *The evidence base of systemic family and couples therapies*. Warrington, UK: Association for Family Therapy.
- Tannen, D. (2007). *You just don't understand: Women and men in conversation*. NY: William Morrow Paperbacks.
- Tilden, T., Gude, T., Sexton, H., Finset, A., & Hoffart, A. (2009). The associations between intensive residential couple therapy and change in a three-source follow-up period. *Contemporary Family Therapy, 32*(1), 69–85. doi:10.1007/s10591-009-9104-8

- Trudel, G., Boyer, R., Villeneuve, V., Anderson, A., Pilon, G., & Bounader, J. (2008). The marital life and aging well program: Effects of a group preventive intervention on the marital and sexual functioning of retired couples. *Sexual and Relationship Therapy, 23*(1), 5–23. doi:[10.1080/14681990701635061](https://doi.org/10.1080/14681990701635061)
- United Nations Economic and Social Affairs Population Division. (2003). *World fertility report*. New York, NY: United Nations.
- Walitzer, K., Dermen, K., Shyhalla, K., & Kubiak, A. (2013). Couple communication among problem drinking males and their spouses: A randomized controlled trial. *Journal of Family Therapy, 35*, 229–251. doi:[10.1111/j.1467-6427.2013.00615](https://doi.org/10.1111/j.1467-6427.2013.00615)
- Weigle, J. B. (2006). *The impact of participating in an Imago therapy workshop on marital satisfaction*. (Unpublished doctoral dissertation). Walden University, Minnesota.
- Whisman, M. A. (2007). Marital distress and DSM-IV psychiatric disorders in a population-based national survey. *Journal of Abnormal Psychology, 116*, 638–643.
- Whisman, M. A., Dixon, A. E., & Johnson, B. (1997). Therapists' perspectives of couple problems and treatment issues in couple therapy. *Journal of Family Psychology, 11*, 361–366. doi:[10.1037/0893-3200.11.3.361](https://doi.org/10.1037/0893-3200.11.3.361)
- Zielinski, J. J. (1999). Discovering Imago relationship therapy. *Psychotherapy, 36*(1), 91–101.