More than One Way to Say I Love You:

An Internet-Based Intervention for Training Flexible Thinking in Romantic Relationships

Karl Cheng-Heng Fua

B.Eng. (Hons.), National University of Singapore, 2003

Ph.D. (Electrical & Computer Engineering), National University of Singapore, 2008

M.A., University of Virginia, 2015

A Dissertation Presented to the

Graduate Faculty of the University of Virginia in

Candidacy for the Degree of Doctor of Philosophy

Department of Psychology

University of Virginia

December 13, 2018

Committee Members:

Bethany A. Teachman, Ph.D. (Chair)

Robert E. Emery, Ph.D.

Joseph P. Allen, Ph.D.

Allison J. Pugh, Ph.D.

Funding Source:

Randy F. Gerson Memorial Grant, American Psychological Foundation

Abstract

Despite the high prevalence of distress in romantic relationships and the significant negative impact of relationship distress on individuals' mental and physical health, a large majority of distressed couples do not seek treatment due to treatment inaccessibility and stigma. Further, there is a high treatment dropout rate for couples who do seek therapy. Hence, there is a need to develop interventions that are accessible and appealing to those experiencing, or at risk for, relationship distress. Cost-effective, internet-based interventions, such as Cognitive Bias Modification (CBM) programs, that target specific cognitive mechanisms to promote flexible thinking, are ideal ways to overcome the barriers to treatment seeking in couples. However, although CBM paradigms have been shown to be effective in diverse domains (e.g., anxiety and depression), these programs have not been applied to distressed couples. In the current project, we developed and pilot-tested a new adaptation of CBM to target psychological flexibility tied to relational problems (i.e., relational flexibility). Specifically, we will: (1) develop a new Aggregated Relational Flexibility measure (ARF); (2) design training materials for a novel CBM program (CBM-FlexC) that targets relational flexibility; and (3) conduct a pilot feasibility and efficacy study of the CBM-FlexC program.

The ARF was developed and validated using four independent samples. In Study 1A, individuals (N = 208) currently in committed relationships (e.g., married or cohabitating for more than 3 months) were recruited online and answered a set of online questionnaires relating to relational flexibility and various aspects of couples' and individual's functioning, to examine the psychometric properties of items shortlisted for the ARF. The new measure was then validated in Study 1B, using a separate online sample (N = 430) of individuals currently in committed relationships, and a subset of this sample (N = 196) was used to establish test-retest reliability. The main results in Studies 1A and IB were replicated in Studies 1C (using an online sample of

260 couples) and 1D (using a laboratory-based sample of 85 couples). The relationship between an individual's relational flexibility and their partner's relationship satisfaction was also examined in Studies 1C and 1D.

The new CBM-FlexC training materials were developed in Study 2, and expert users (N = 4) and end-point users (N = 7) were recruited to provide feedback (e.g., on clarity and relevance of training materials) via qualitative interviews. Finally, the feasibility and efficacy of the CBM-FlexC program were evaluated in Study 3, using an online sample of currently distressed couples (N = 18). Using a multiple baseline design, participants underwent at least 3 baseline assessment sessions, followed by 6 online sessions of CBM-FlexC training over two weeks, followed by a one-month follow-up. Relational flexibility was assessed after the second, fourth, and sixth training sessions. Additionally, participants completed online questionnaires pre- and post-intervention, and also at the 1-month follow-up. It was hypothesized that couples' relational flexibility would improve from baseline to post-training and at the 1-month follow-up, and that the increase in an individual's relational flexibility would predict an increase in both their own and their partner's relationship satisfaction. Overall, as expected, CBM-FlexC training resulted in higher relational flexibility and relationship satisfaction compared to at baseline, and increases in an individual's relational flexibility significantly predicted increases in their relationship satisfaction. Additionally, these improvements were maintained one-month after training. However, contrary to expectations, increases in an individual's relational flexibility were not significantly associated with changes in their partner's relationship satisfaction.

The current project paves the way for a larger-scale internet-based intervention study that targets distressed couples with limited access to more traditional tools and resources (e.g., couples therapy) due to lack of resources or stigma. This is the first study, to our knowledge, to apply the CBM paradigm to training flexible thinking in couples. Once validated in larger trials,

this new intervention has the potential for wide dissemination as an evidence-based tool for distressed couples to help resolve their relationship problems.

To my family, thank you for your

unconditional love and support.

I love you all.

Acknowledgements

What an adventure this has been! Looking back on the long, somewhat circuitous road I took to find this line of work I truly believe in and love, I am grateful for the warmth, love, encouragement, and support from all those who accompanied me on this journey.

First of all, I want to thank my advisor, Bethany Teachman. Bethany, I have told you many times before, and I will say it again: Thank you for believing in me and giving me the chance to pursue a career that I love and find fulfilment in every single day. You, quite literally, changed my life, and words cannot describe how grateful I am for that. People thought I was "crazy" for my drastic, and somewhat late, career switch, but now, I am more certain than ever before that it was the right choice. Thank you for being such an awesome mentor and friend. Your commitment to doing great research and clinical work is truly inspiring, and your positivity and cheerfulness made it such fun to work with you. So, thank you, for being a great role-model, for your support and mentorship, for your warmth, optimism, and good cheer, and for welcoming me into the PACT family.

I want to thank the PACT family for being my family away from home. It has been such a privilege and joy to work alongside all of you. To Shari Steinman, Jen Green, Meg Reuland, Ann Lambert, Jessica Beadel, Gena Gorlin, Jeff Glen, and Alex Werntz, for welcoming me into the lab and showing me the ropes when I was a wee-lil'-fledgling PACT member. I want to thank Nauder Namaky and Diheng Zhang for your friendship, stimulating conversations, and countless lunch hours of fun and gossip--you have added more joy and laughter to my days in Gilmer than you can imagine. Thank you to Miranda Beltzer (and your "tiny hands"!), Katharine Daniel, Jeremy Eberle, Henry Behan, Philip Chow, Alex Daros, Julie Ji, Nicola Hohensee, and Claudia Calicho-Mamani for being such awesome labmates and bringing me such joy every time I see

you. All of you, past and present labmates, have shaped (in very positive ways!) so much of my life in the past half-a-decade, and brought me so much wacky laughter and given me so much support. It is hard to put it all into words, but I appreciate all of you--very much--and I look forward to more years of reminiscing and laughter at lab reunions!

Thank you to my dissertation committee members, Bob Emery, Joe Allen, and Allison Pugh, for your great insights, helpful feedback, and generous moral support on the project. The project is much better now because of you. Thank you also to all the research assistants I have had the pleasure of working with on so many projects: Sekar Novika, Somil Chugh, Tairuo Ge, Virginia Clemo, Emily Geyer, Sarai Arbus, Aradhana Khubani, Kirby Eule, Claire Friedman, and Maryam Farooqui.

To my wonderful, wonderful cohort-mates, Rachel Narr and Matthew Domiteaux: You cannot imagine how much you both mean to me. I love you both. Thank you for the moral support, smart banter, gossip, and laughter, these past five years that pulled us through Assessment, TEAM, practicums, supervision hours, internship applications/interviews, crumbling lunatic asylums, and the many more challenges we surmounted together. I enjoyed every single minute with you two, and I look forward to creating more memories with you and having our annual cohort retreats together.

Last, but definitely not the least, I want to thank my family for the unconditional support, love, and trust, freely and generously given to me all my life. I will not be who I am today without all of you. To my parents, Sinkan Fua and Toysiew Koh: Ba, Ma, thank you for being the awesome, loving parents you are, for always being open and supportive of my choices and pursuits, for believing in me, and for being the constant, stable pillars of support I know I can always rely on and come back to any day. Every single day, I long to hear your gentle voices and I miss your loving smiles. Ba, I miss you a lot, and I hold dearly in my heart your gentle

words of encouragement and kindness, your smile, your strength, your optimism, and your mindful approach to life. To Minghuey Fua, Yinghuey Fua, and Zhengqi Pan: Jeleng, Jewei, and Qi, your care, love, concern, and support means the world to me. I miss you all and look forward to many more years of crazy laughter and fun together. Finally, I want to thank my wonderful partner, Garland Easter III, for being my loving and reliable companion. I cannot imagine what my life would have been like these past few years without you by my side. You have brought love, stability, and fulfillment into my life, and you are the rock that I can rely on and trust. I look forward to walking with you, hand in hand, as we undertake the exciting journey that is life.

Table of Contents

Abstract ii

Acknowledgements vi

Table of Contents ix

Training Flexible Thinking in Relationships 1

Relational Inflexibility and Relationship Distress 2

Promoting Relational Flexibility via Couples Interventions 5

Online Interventions 6

Overview and Hypotheses 10

Study 1: Measuring Relational Flexibility 11

Methods 12

Participants and Recruitment 12

Recruitment/Sampling Considerations 15

Materials 16

Procedure 18

Plan for Analyses 20

Study 1A Results 21

Exploratory Factor Analyses 21

Validity of the ARF measure 22

Convergent and Divergent Validity....22
Incremental Validity....22
Study 1B Results....23
Confirmatory Factor Analysis....23
Item Reduction....24
ARF Scale Validity....24
Construct Validity....24
Incremental Validity....25
Test-Retest Reliability....25
Study 1C and Study 1D Results....25
Association between Relational Flexibility and Relationship Satisfaction....25
Individual's Relational Flexibility and Partner's Relationship Satisfaction....26
Summary of Study 1 Results....26
Study 2: Piloting Training Materials for Novel CBM-FlexC Paradigm....27
Methods....27
Participants....27
Materials....27
Procedure....30
Study 2 Results....31
Findings from Qualitative Interviews....31

Summary of Study 2 Results....34

Study 3: Feasibility and preliminary test of efficacy of the CBM-FlexC paradigm 34

Methods....34

Participants 34

Materials 36

Procedure....38

Hypotheses and Plan for Analyses 40

Results....42

Participant Adherence and Participant Characteristics 42

Baseline Stability....43

Quantitative Analyses to Assess Change Over Time....44

Changes in Relational Flexibility....44

Changes in Relationship Satisfaction 45

Changes in General Cognitive Flexibility 45

Changes in Differentiation of Self 46

Exploratory Analyses: Does Sex Matter? 47

Exploratory Analyses: How does an individual's relational flexibility impact their partner's relationship satisfaction? 47

Graphical Analyses 48

Qualitative Analysis of Exit Interviews 50

Summary of Study 3 Results....55

Discussion....55

Composition of Relational Flexibility....57

General vs. Relational Flexibility....58

Overall Sample versus Individual Change....59

Changes in Relational Flexibility over Time and Number of Intervention Sessions....60

Changes and Stability of Relationship Satisfaction....62

Impact of Relational Flexibility on Partner's Relationship Satisfaction....63

Possible Mechanisms Linking the Intervention to Relationship Outcomes....65

Technological Considerations....67

Limitations and Conclusions....68

References....70

Tables....79

Table 1. *Factor loadings of 25 shortlisted Aggregated Relational Flexibility (ARF) items on the factors Perspective Taking (PT), Negative Reactions (NR), and Rigidity (RI).* 79

Table 2. *Correlation matrix for exploratory factor analysis (Study 1A)*....81

Table 3. *Results from model comparisons examining incremental validity of ARF in predicting couples satisfaction and negative relationship attributions (Study 1A)*....82

Table 4. *Correlation matrix for confirmatory factor analysis (Study 1B)*....83

Table 5. *Results from model comparisons examining incremental validity of ARF (8-item version) in predicting couples satisfaction (Study 1B)* 84

Table 6. *Correlation matrix of ARF full scale and subscale scores at times 1 and 2 (Study 1B)* 85

Table 7. *Correlations between an individual's relational flexibility and their partner's relational flexibility and relationship satisfaction* 86

Table 8. *Means and Standard deviations on measures in Study 1* 87

Table 9. *Means and standard deviations of measures in Study 3* 88

Figures 89

Figure 1. Structural Equation Modeling (SEM) model confirming the factor structure of the ARF (14 items). 89

Figure 2. Structural Equation Modeling (SEM) model of the 8-item ARF. 90

Figure 3. Change in mean levels of relational flexibility over sessions. 91

Figure 4. Change in mean levels of relationship satisfaction over sessions. 92

Figure 5. Change in mean levels of general cognitive flexibility over sessions 93

Figure 6. Changes in mean levels of the 4 differentiation of self subscales over sessions. 94

Figure 7. Changes in partner's relationship satisfaction as a function of session and an individuals' relational flexibility. 95

Figure 8. Sample plots of participant relational flexibility over time. 96

Appendix A. Sample Training Materials 97

Appendix B. Example of CBM-FlexC Training Materials Assessment Protocol 103

Appendix C. Sample Training Schedule....104

Appendix D. Measurement Plan....105

Appendix E. Questionnaires....107

Appendix G. Qualitative Exit Interview....120

Training Flexible Thinking in Relationships

Approximately 50% of married couples file for divorce, and half of these divorces occur within the first seven years of a marriage (Kreider & Fields, 2002). Distress in romantic relationships has a profound impact on individual mental health and is highly correlated with the occurrence of mood, anxiety, and substance use disorders (Whisman & Uebelacker, 2006), domestic violence (Cascardi & Vivian, 1995), distress in broader social relationships (e.g., with extended family, friends), poorer general health, and reduced responsiveness to treatment of mental health issues (O'Farrell et al., 1998; Whisman, 2001). Despite the high costs to individuals and society, only about 20% of distressed couples seek help in the form of couples therapy (Johnson et al., 2002). Furthermore, there is a high dropout rate for those who do seek help (Masi et al., 2003). Given the alarming rate of divorce and serious personal and economic impact of relationship distress, we need to develop effective interventions that target key cognitive mechanisms that underlie dysfunctional relationships, and that are accessible and appealing to those experiencing, or at risk for, relationship distress.

In the current project, we developed and validated the Aggregated Relational Flexibility (ARF) measure, an instrument that measures relational flexibility (i.e., psychological flexibility in relationships), an important but understudied cognitive construct in the context of relational functioning. We then designed and piloted CBM-FlexC, a new adaptation of the Cognitive Bias Modification paradigm to target relational flexibility. Once validated in larger trials, this new intervention has the potential for wide dissemination as an evidence-based tool for distressed couples to help resolve their relationship problems.

Relational Inflexibility and Relationship Distress

Psychological inflexibility is characterized by overly rigid styles of thinking in which an individual engages in perseverative thinking patterns and inflexible responses that reduce the ability to engage in versatile coping or problem solving (Kashdan & Rottenberg, 2010). It is an established, key cognitive mechanism that drives individual psychological well-being (Kashdan & Rottenberg, 2010), and is a major risk factor for psychopathologies, such as anxiety (Thayer et al., 1996) and depression (Rottenberg et al., 2005). Likewise, in the domain of relationship functioning, we use the term *relational inflexibility to refer to rigid styles of thinking and emotional expression and experience that function as core mechanisms that drive many maladaptive behaviors in couples.* These behaviors directly impact relational well-being, so relational inflexibility has been hypothesized to be an important feature of unhealthy functioning in couples (Kelly et al., 2003). While related to psychological flexibility, relational flexibility involves aspects specific to relationship functioning and dyadic social interactions (e.g., taking the perspective of another, reacting to interpersonal conflict), and thus provides a more nuanced characterization of the role flexibility plays in relationships. Despite the seemingly critical role relational flexibility plays in relationship satisfaction, to the best of our knowledge, relational flexibility has not been explicitly studied as a construct. However, there is substantial research on separate but core aspects of relational flexibility. These core features can be organized into two groups--the cognitive and affective components--that together form the aggregated construct of relational flexibility.

The *cognitive component of relational inflexibility* involves overly rigid and unchanging beliefs about partners and relationship events. Specifically, these rigid beliefs typically involve why relationship events occur (e.g., a belief that a partner had purposefully brought about a negative relationship event), how relationship events should or should not occur, how partners

should or should not behave, the likelihood of future negative events (Baucom et al., 1989), and the degree to which an aspect of a partner or the relationship is amenable to change or improvement (e.g., that a negative partner behavior is never going to change; Fincham & Bradley, 1992). These rigid beliefs impair individuals' ability to adopt and understand the perspective of their partners emotionally and cognitively, leading to lower relationship satisfaction (Peloquin & Lafontaine, 2010).

Relational inflexibility is likely bidirectionally related to the negative cognitive biases individuals experience in the context of their relationships. These cognitive biases include attention (fast attentional shift and delayed disengagement from negative events that occur in the relationship; Sillars, et al., 2000), interpretation (interpreting ambiguous partner behaviors, such as a partner being quiet, negatively; Finn et al., 2013), and memory biases (bias toward remembering more negative relationship events; Halford et al., 2002). Cognitive biases can also include negatively biased expectancies and attributions for relationship-specific events (Baucom, Epstein, & LaTaillade, 2002), and believes that partners are unwilling, or incapable, of changing problematic behavior (Fincham & Bradley, 1992). We expect that the selective processing of relationship-relevant cues that occurs with these cognitive biases helps maintain relational inflexibility. Moreover, difficulty taking a partner's perspective and responding flexibly likely decreases opportunities to disconfirm cognitive biases, resulting in maladaptive interaction patterns that lead to negative consequences for the maintenance of healthy relationships.

The *emotional component of relational inflexibility* –that is, rigid ways of responding to, experiencing, and expressing emotions in relationship contexts– is likely closely connected to an individual's ability to adequately understand and process the broad range of emotions felt during relationship-specific events. This inflexibility and lack of understanding appears to be especially damaging to relationships given perseverative emotional experiences focus on and

intensify unproductive emotions (Snyder et al., 2006). The presence of excessive negative emotions produces maladaptive patterns of interactions and intense, negative emotional reactions to relationship events (e.g., demand-withdraw behavior; Johnson, 2005; or high emotional reactivity; Skowron, 2000), leading to lower satisfaction in couples (Gottman et al., 1998; Huston et al., 2001). Additionally, distressed relationships are characterized by higher rates and longer duration of experienced and expressed negative affect, which leads to higher levels of reciprocal expression of negativity to partners and inversely lower levels of positive affect (Snyder et al., 2005). Relationship satisfaction (and likelihood of divorce) is further affected by rigid perceptions of partner affect and beliefs about what might be the cause of an expressed negative affect by a partner.

Hence, directly promoting relational flexibility in couples could potentially improve the quality of relationships and relationship satisfaction via a number of cognitive and affective pathways. For example, increasing relational flexibility could: improve an individual's ability to relate to and communicate with their partner (e.g., by being able to take the perspective of the other person); increase positive and more benign attributions and interpretations individuals make of their partner's affect and behavior; reduce intensity of emotional reactivity to relationship events; enhance empathy within relationships; and improve understanding of the relationship between one’s own affect/thoughts/behavior and their partner's affect/thoughts/behavior (Cook et al., 1995). Given the important role of relational flexibility in couples functioning, it is important to develop reliable tools for assessing relational-specific psychological flexibility, and efficacious interventions that can directly promote relational flexibility in couples.

Promoting Relational Flexibility via Couples Interventions

Various forms of empirically supported couples therapy have been shown to be effective in treating relationship distress (Christensen & Heavey, 1999), with effect sizes generally ranging from medium to large, and with an overall mean effect size of approximately 0.84 (Shadish & Baldwin, 2003). In general, successful approaches tend to involve multi-faceted approaches that involve some combination of changing and reframing cognitions and attributions, fostering emotional acceptance by focusing on deeper and more flexible exploration of emotions, and consideration and adoption of different positive behavioral strategies (Snyder et al., 2006) – all key aspects of the relational flexibility construct. For instance, Behavioral Couples Therapy focuses on improving communication and changing behavioral styles, while Integrative Behavioral Couples Therapy includes the fostering of emotional acceptance in addition to communication and behavioral strategies; Cognitive Behavioral Couples Therapy focuses on identifying and addressing cognitive distortions and biased interpretations individuals have with respect to their partners' expressed emotions and behaviors, in addition to building behavioral change strategies; Integrative Systemic Couples Therapy focuses on reframing and changing meaning attributed to unproductive and negative patterns of interactions; and Emotion Focused Couples Therapy focuses on improving emotional awareness and regulation (Greenberg & Pascual-Leone, 2006).

Current empirically supported couples therapies, though efficacious, require couples to attend, in-person, regular therapy sessions (median of 12 sessions; Doherty & Simmons, 1996) with appropriately trained therapists or counselors. Hence, these therapies have limited reach in terms of their availability and accessibility to distressed couples, and do little to mitigate the stigma associated with in-person help-seeking behavior (even when help is available). Furthermore, while the absolute gains (post vs. pre-treatment) from receiving couples therapy

tend to be robust over time (e.g., Christensen et al., 2006; Christensen et al., 2010), evidence suggests that the effect sizes of treatment gains tended to reduce significantly over time (e.g., see Christensen & Heavey, 1999). Thus, it is important that treatments continue to be readily accessible, in various forms, by individuals and couples to help couples maintain the gains they have made over time.

Online Interventions

In recent years, there has been a push by clinicians and researchers to exploit the wide reach of the internet to disseminate treatments to high-risk populations with low-treatment seeking behavior. Internet-based interventions are potentially low cost, and easily accessible and re-accessible. Despite increasing evidence indicating that online adaptations of evidence-based treatments for individuals are efficacious and comparable to traditional face-to-face treatment (e.g., Cuijpers et al., 2008), there are limited online versions of evidence-based treatments available for couples, and the majority of online resources for couples are in the form of psycho-education (e.g., online resources provided by the Gottman Institute; www.gottman.com) or relationship forums (e.g., relationship advice subreddits on Reddit.com; www.talkaboutmarriage.com, as discussed by Georgia and Doss, 2013). Nonetheless, the few online evidence-based interventions currently available for couples have been shown to be effective in preventing relationship distress (e.g., ePREP; Braithwaite & Fincham, 2007; Georgia & Doss, 2013). Specific to couples distress, an online-CBT program, based on Integrative Behavioral Couples Therapy (IBCT; Christensen et al., 2010), developed by Doss et al. (2013) is currently being validated with very encouraging outcomes (Doss et al., 2016). This 8-hour program includes relationship education, detailed assessment, and feedback regarding areas of difficulties specific to a couple. The program guides distressed couples through online cognitive restructuring exercises that are tailored to the couple's issues, and couples learn strategies for

improving communication and acceptance. Couples report significant improvements in multiple outcome measures including relationship satisfaction, perceived relationship quality, and even perceived work functioning and quality of life (Doss et al., 2016). In general, online CBT programs tend to be similar to in-person CBT-based treatments, in that they target various important aspects of relationships (e.g., acceptance, attributions, communication styles/techniques). These programs rely on the readiness of couples to discuss and confront key relationship challenges, and be able to commit to a relatively long treatment program. However, a majority of distressed couples may not be ready to do so, based on the low rate of treatment-seeking behavior (Doss et al., 2009). Therefore, it is important to explore new internet-based, empirically-supported interventions that target high-risk couples in ways that may be less intimidating and that reduce the stigma surrounding therapy.

An ideal way to reach out to high-risk couples experiencing multiple barriers to seeking treatment is through the use of cost-effective, internet-based Cognitive Bias Modification (CBM) programs (Mathews & MacLeod, 2012). CBM uses basic learning principles to shift specific cognitive biases in many domains (for example, anxiety and depression; see MacLeod & Mathews, 2012, for review). Thus, these programs do not require individuals to explicitly confess/discuss their personal relationship challenges; a step they might not be ready to take. Additionally, CBM programs are usually short and game-like computerized tools that are easily portable in an internet-based format. The number of sessions required for CBM training is variable in the literature, typically ranging between 1 to 8 for interpretation bias training (e.g., Beard, 2011; Hakamata et al., 2010), though the length of each session is typically short (Beard, 2011), thus requiring much less time commitment from couples than standard face-to-face treatment. Additionally, the high portability of CBM treatments over the internet means that these treatments can be disseminated at very low cost. Thus, CBM treatments have a

potentially wide-reach, making them ideal early interventions to address stigma, accessibility, and low-treatment seeking behavior in high-risk couples. Furthermore, low cost CBM treatments would expose couples to the benefits of evidence-based treatments, and could serve as an excellent stepping stone for couples to seek further, more intensive evidence-based treatment (e.g., face-to-face couples therapy) if needed. Additionally, CBM sessions can be completed independently by each partner in a couple, thus reducing the time barrier as couples do not have to coordinate and invest a fixed, commonly available time to completing the intervention sessions. Though CBM sessions are typically completed independently by each individual, these interventions may have the potential of improving relationships even when completed by only one partner. For instance, an individual who is able to flexibly adopt the perspective of their partner (after receiving individual CBM training) might behave more positively in their relationship, which may then encourage reciprocal positive behaviors from their partner and trigger a recursive dynamic in the relationship that leads to improvement in relationship satisfaction over time (Marigold, Holmes, and Ross, 2010).

CBM programs follow multiple formats, depending on the specific cognitive mechanism targeted. In particular, CBM-Interpretation bias modification programs (CBM-I) target anxious individuals' tendency to selectively interpret ambiguous situations in negative ways, typically by training individuals to assign more benign interpretations of ambiguous situations. In one widely used CBM-I paradigm, training is achieved by having individuals practice assigning benign or positive interpretations to a series of fictional ambiguous scenarios. CBM-I has demonstrated efficacy in varied domains, such as anxiety (Salemink et al., 2009; Steinman & Teachman, 2010), depression (Wells & Beevers, 2010), and addiction (Fadardi & Cox, 2009). The training paradigm has also been successfully adapted to change appraisal styles by altering negative appraisals of intrusive thoughts typically seen in depressed individuals (Lang et al., 2009).

It is important to note that results examining the efficacy of CBM training paradigms, especially those conducted online, are mixed (see Cristea et al., 2015, for meta analyses). One reason may be that most current CBM programs emphasize positivity (e.g., consistently training positive interpretations of ambiguous situations). Though increasing positive interpretations potentially addresses the negative biases individuals (especially those who are distressed) experience, we expect that the more critical cognitive change to promote is flexibility in thinking because this allows individuals to expand their repertoire of responses so they can more readily respond to dynamically changing situational demands. Given that counter-regulation, where biases dynamically change to allocate attention to valenced information that is opposite to the current affective-motivational state, is an important mechanism in adaptive information processing (Rothermund, Voss, & Wentura, 2008), rigid, unchanging biases (in either positive or negative directions) can be problematic. For example, if one's partner came home from work and engaged in what might be perceived as a rude behavior (e.g., slamming the door), a strategy to think only about the positive aspect of the situation might not be sufficient or suitable, because positive aspects might be difficult to identify or relate to in the situation, and a rigid positivity focus might even perpetuate unhealthy relationship behaviors (e.g., passivity and inaction in problematic relationship situations). In this example, flexible thinking in terms of examining both positive and negative potential reasons for his/her behavior or considering the different ways oneself or one's partner might be thinking and feeling, can lead to a more balanced and realistic understanding of the situation and increased empathy, which can promote productive conversations and problem solving, effectively regulate emotional reactions, and reduce the chance of high intensity negative reactions and confrontations. Thus, it is important to develop novel adaptations of CBM programs that go beyond rigidly training positivity in isolation, and instead target *flexibility*, in this case relational flexibility, more directly.

At the present time, the role of CBM interventions for couples in the context of other couples interventions (e.g., traditional in-person couples therapy) remains unclear. CBM interventions might be the only interventions that can be accessed by certain populations (e.g., those in rural areas, or those facing other cost/accessibility constraints), or might be a stepping-stone to or used in conjunction with other treatment modalities, or might be efficacious only for certain subsets of the population (e.g., younger couples) and not others. However, by being targeted, simple, brief, and cost effective, CBM interventions address many of the barriers to treatment and have the potential to be an excellent resource for distressed couples. Thus, while the current project is not designed to test the interesting question of whether and when CBM will be sufficient as a stand-alone treatment and when it will be most helpful as an adjunct to other types of care, we are excited by its potential for wide dissemination and integration with different models of care.

Overview and Hypotheses

In this research, we developed a novel application of the CBM paradigm to train and improve relational flexibility in distressed couples (CBM-FlexC). To the best of our knowledge, this is the first attempt to adapt the CBM paradigm to target relational flexibility, a key cognitive mechanism implicated in relationship distress, and the first attempt to apply CBM to help couples experiencing relational distress. In the first study, we developed an Aggregated Relational Flexibility (ARF) measure by drawing upon and adapting existing instruments that either measure general psychological flexibility or separate aspects of the construct (e.g., perspective taking in a relationship, empathy, openness to considering alternative viewpoints). Currently available measures tend to focus on isolated components of relational flexibility; it is thus important to design an aggregated measure that can account for the combined effect of these components. The novel relational flexibility measure was used as the primary outcome

measure for evaluating the efficacy of the proposed CBM intervention for couples. In the second study, the CBM-FlexC training materials were developed, and an expert panel of couples researchers and a separate group of couples from the community were recruited to provide feedback on the clarity, relevance, and appeal of the training materials and program using qualitative interviews. Finally, in the third study, a pilot feasibility and efficacy study of the CBM-FlexC program was conducted. For the pilot study, a multiple baseline study design (Barlow & Hersen, 1984) was used to allow efficient, initial assessment of the causal link between the CBM-FlexC intervention and improvements in relational flexibility and relationship satisfaction. Distressed couples were recruited online. These couples underwent a series of baseline assessments, followed by six web-based sessions of CBM-FlexC (two sessions per week, over three weeks), and a one-month follow-up assessment. The pilot study provided important preliminary information about ways the CBM-FlexC program can be improved, and the feasibility and likely efficacy of applying the program in a future larger-scale study. Given CBM's prior positive outcomes in other domains (see e.g., Hallion & Ruscio, 2011), it was expected that couples' relational flexibility, as measured by the new ARF, will be higher post-training and at the one-month follow-up compared to levels measured at baseline. It was further hypothesized that the extent to which an individual's relational flexibility is increased will be positively associated with degree of improvement in that individual's and their partner's relationship satisfaction.

Study 1: Measuring Relational Flexibility

While measures of general psychological flexibility and various aspects of relational flexibility exist, there is as yet no direct instrument for assessing relational flexibility in couples. In the first study, we developed a reliable, valid measure of relational flexibility. Such a measure

is necessary to evaluate if the novel CBM-FlexC program shifts relational flexibility as anticipated. In Study 1A, items considered for the Aggregated Relational Flexibility (ARF) measure, along with existing measures of psychological flexibility and relationship functioning, were administered to an online sample to identify those items from the ARF that are most associated with relationship distress, and that have good psychometric properties (based on an exploratory factor analysis and evaluation of internal consistency). The selected items were then validated in Study 1B using a confirmatory factor analysis with a separate online sample. Test-retest validity was established by administering the aggregated relational flexibility measure twice, approximately 1-2 weeks apart, to a sub-sample of those recruited in Study 1B. In Studies 1C and 1D, the final set of ARF items were administered to two independent samples of couples who were recruited as part of two larger, unrelated studies conducted at two different sites. Studies 1C and 1D provided additional validation of the ARF, and allowed for the examination of the association between an individual's relational flexibility and their partner's relationship satisfaction.

Methods

Participants and Recruitment

Study 1A. An internet-based sample of English speaking individuals living in the United States and who are currently in a monogamous romantic relationship for at least 3 months (following Gonzaga et al., 2007) was recruited on Amazon Mechanical Turk (MTurk) to participate in the study. Out of the 338 individuals who responded to the screener on MTurk (MTurk), 208 individuals (Female = 52.88%; Male = 46.64%; Others = .48%) qualified for the study and were administered the full battery of questionnaires. Participants were 22 to 69 years old (M = 40.2, SD = 10.6), and reported their race as 77% White, 7% Black, 12% Asian, 1%

Native American/Alaskan Native, 1% Hawaiian/Pacific Islander, and 2% others/unknown. The majority of participants identified as non-Hispanic/Latino (91%). All participants reported that they were in a romantic relationship, and 69% of these participants reported that they were also currently engaged, married, or in civil union or domestic partnership. All participants reported having been in their current romantic relationship for more than 3 months, and 78% reported having been in their current relationship for more than 3 years. Most participants (96%) reported that they were in heterosexual relationships. Participants were paid $1.50 for completing the study.

Study 1B. Study 1B comprised 2 parts. In Part 1, a separate internet-based sample from Project Implicit was screened (total screened = 889). Those individuals (N = 430; Female = 61.63%, Male = 38.37%) who met the same inclusion criteria as Study 1A were recruited for Study 1B. Participants were 19 to 87 years old (M = 43.23, SD = 13.5), and reported their race as 79% White, 8% Black, 3% Asian, 1% Native American/Alaskan Native, 6% mixed, and 2% others/unknown. The majority of participants (81%) identified as non-Hispanic/Latino. All participants reported that they were in a romantic relationship, and 82% of these participants reported that they were also currently engaged, married, or in a civil union or domestic partnership. All participants were in relationships longer than 3 months, and 87% of these participants reported having been in their current relationship for more than 3 years. The majority of participants (90%) were in heterosexual relationships. Overall, the sample was similar to the MTurk sample used in Study 1A.

Of the 430 participants who completed Part 1 of Study 1B, a sub-sample of 196 participants also participated in Part 2, which was a retest administered approximately 1 week after participants completed Part 1. The demographic characteristics of the sub-sample were similar to that of the main sample (used in Part 1) in terms of sex (Female = 66.84%, Male =

33.16%), age (M = 44.43, SD = 14), race and ethnicity (89% White, 4% Black, 3% Asian, .5% Native American/Alaskan Native, 3% mixed, .5% others/unknown), and relationship characteristics (80% were currently engaged, married, or in civil union or domestic partnership, 84% have been in their current relationship for more than 3 years; 88% were in heterosexual relationships). Participants who completed both parts of Study 1B were entered in a lottery for a $100 Amazon gift card; those who only completed Part 1 of Study 1B were not paid.

Study 1C. Participants were from an internet-based sample (N = 260 couples) recruited by researchers at the University of Miami for an online couples intervention based on procedures outlined in Doss et al. (2016). These participants completed the ARF as part of the battery of questionnaires administered at the beginning of the larger study. The majority of individuals (42%) in the sample were between 25-34 years old, and 31% were between 35-44 years old. Participants reported their race as 58% White, 25% Black, 2% Asian, 1% Native American/Alaskan Native, .4% native Hawaiian or Pacific Islander, 6% mixed, 8% others/unknown, and 88% identified as non-Hispanic/Latino. All participants were currently in a relationship with the partner they completed the questionnaire with, and had been in the relationship for at least 3 months (M = 5.8 years, SD = 5.09 years).

Study 1D. Participants were community couples who were part of an ongoing, laboratory-based longitudinal study on marital functioning at the University of Arizona. Participants completed the ARF as part of the battery of questionnaires administered at the beginning of the second wave of data collection for the larger study. A total of N = 85 heterosexual, married couples completed the ARF. These participants were between 19 to 49 years old (M = 28.4, SD = 5.9), and reported their race as 24% White, 0.6% Black, 1% Asian, 45% Hispanic/Latino, 0.6% Native American/Alaskan Native, 0.6% Native Hawaiian or Pacific

Islander, 27% mixed, 1% others/unknown. Notably, the majority (71%) of the participants identified as being Hispanic/Latino.

Recruitment/Sampling Considerations. Given that there are no significant differences in most relationship challenges experienced by heterosexual vs. same-sex couples (Kurdek, 2004), recruitment for all studies (except samples in Study 1C that were recruited as part of two other studies) did not discriminate between types of couples. Internet-based sampling permitted for a relatively broad sampling of the population at a low cost. Further, considering the demographics of typical participants on internet-based recruitment platforms (e.g., on Amazon Mechanical Turk, where approximately 68.4% are currently in a romantic relationship and 48.1% are currently married; Shapiro et al., 2013), the proposed recruitment criteria (which follows Gonzaga et al., 2007) maximized the pool of participants.

There is no consistent recommendation in the literature for minimum sample sizes and subject-to-variable ratios for conducting exploratory factor analyses--suggested sample sizes range from N=100 (Gorsuch, 1983) to N=500 (Comrey & Lee, 1992), and subjects-to-variable ratios ranged from 2:1 (Kline, 1979) to 20:1 (Hair et al., 1995). Multiple research on the minimum subject-to-variable ratio required to recover the population factor structure appear to indicate that a subject-to-variable ratio of 3 - 3.9 might be sufficient (e.g., MacCallum et al., 1999; Arrindell & van der Ende, 1985). Given the lack of consistent guidelines, we elected to follow Bryant and Yarnold's (1995) recommended 5:1 subjects-to-variable ratio. Our sample sizes in Study 1A (N = 208) and Study 1B (N = 430) provided adequate to excellent subjects-to-variable ratios of 8:1 (with 25 shortlisted items in Study 1A) and 31:1 (with 14 items in Study 1B, retained after Study 1A), and satisfied the chosen criteria. Additionally, given that a ratio of 10 participants per item is recommended for the purposes of establishing test-retest validity (e.g., Wolf et al., 2013), in Study 1B, we recruited until at least 140 participants completed the

measure twice (i.e., 10 times 14, the number of items from the ARF retained after Study 1A). The final number of participants who completed both parts of Study 1B was 196.

Materials

Online Pre-screening survey. Online MTurk participants answered a brief screener consisting of three questions about their relationship status, length of relationship, and if the relationship was long-distance. See Appendix E for the items.

Demographics information. Participants provided information about their age, gender, race/ethnicity, citizenship, education, marital status, employment status, and living situation.

Relational Flexibility. The Aggregated Relational Flexibility (ARF) measure comprises 25 shortlisted items (11 reverse scored) adapted from existing measures, and measures individuals' flexibility in relationships. It uses a 5-point Likert scale (where 1="Not at all characteristic of me", and 5="Extremely characteristic of me").

General Psychological Flexibility. The Cognitive Flexibility Inventory (CFI-20; Dennis & Vander Wal, 2010) is a 20-item self-report questionnaire that measures individuals' tendencies to think flexibly when they encounter difficult situations (e.g., "When in difficult situations, I consider multiple options before deciding how to behave."), using a 7-point Likert scale (where 1="Strongly Disagree", and 7="Strongly Agree). The CFI-20 comprises 2 subscales related to individuals' perceived ability to perceive and generate alternative perspectives and solutions to situations (CFI-Alternatives subscale), and the degree of control individuals perceive themselves as having in these difficult situations (CFI-Control subscale). Due to practical limitations and to maintain participant engagement, a shortened 10-item version of CFI, CFI-10, was used in Studies 1B and 3. The CFI-10 comprises items that had factor loadings greater than .70 (based on statistics reported in Dennis & Vander Wal, 2010), resulting in 6-items from the CFI-Alternatives subscale and 4-items from the CFI-Control subscale (see

Appendix E for items). Results from Study 1A showed that correlations between the shortened CFI-10 and other administered measures were similar to those obtained using the full CFI-20 scale (see Table 2). Unless otherwise indicated, CFI refers to the CFI-10 for the remainder of the paper. The CFI was used to establish convergent validity, and as a control variable to establish incremental validity in predicting indices of general couple's satisfaction.

Aspects of Relational Functioning. The Interpersonal Reactivity Index for Couples (IRIC; Peloquin & Lafontaine, 2010) is a 13-item self-report questionnaire that measures empathy and perspective-taking in relationships (e.g., "I try to look at my partner's side of the disagreement before I make a decision."), using a 5-point Likert scale (where 0="Does not describe me well", and 4="Describes me very well"). The IRIC comprises 2 subscales related to empathic concerns and perspective taking. The perspective-taking subscale of the IRIC was used to establish convergent validity, and as a control variable to establish incremental validity in predicting indices of general couple's satisfaction.

The short Relationship Attribution Measure (RAM; Fincham & Bradbury, 1992) is a 4-item measure of different attributions individuals make in relationships. Each item comprises a statement describing a hypothetical negative partner behavior, followed by 7 statements each related to a type of attribution (e.g., for causal (locus) attributions, "My partner's behavior was due to something about him/her") rated on a 6-point Likert scale (where 1="Disagree Strongly", 6="Agree Strongly"). The RAM was used to establish predictive validity (i.e., those with lower relational flexibility were expected to make more negative causal attributions about their partner's negative behaviors and place more responsibility and blame on their partner).

The short Couples Satisfaction Index (CSI-4; Funk & Rogge, 2007) is a 4-item self-report questionnaire that measures general relationship satisfaction in couples. The CSI-4 comprises 1 item about "the degree of happiness, all things considered, of [the] relationship" (using a 7-point

Likert scale; where 0="Extremely Unhappy", and 6="Perfect"), and 3 items about aspects of relationship satisfaction (e.g., "I have a warm and comfortable relationship with my partner") rated on 6-point Likert scales (e.g., 0="Not at all true", and 5="Completely true"). The CSI-4 was used to establish predictive validity (i.e., those with lower relational flexibility were expected to have lower relationship satisfaction).

Depression and Anxiety Symptoms. The Depression, Anxiety, and Stress Scales (DASS; Lovibond & Lovibond, 1995) is a 21-item measure of recent symptoms related to depression, anxiety, and stress (e.g., "I felt I was close to panic."), using a 4-point Likert scale (where 0="Did not apply to me at all", and 3="Applied to me very much, or most of the time"). The Patient Health Questionnaire for Depression and Anxiety (PHQ-4; Korenke et al., 2009) is a brief 4-item measure of depression and anxiety symptoms (e.g., "Feeling nervous, anxious, or on edge") using a 4-point Likert scale (where 0="Not at all", and 3="Nearly every day"). The DASS and PHQ-4 were used to establish discriminant validity (i.e., the measure of relational flexibility was expected to have lower relations with measures of anxiety and depressive symptoms than it does with the measures of cognitive flexibility and perspective taking, though some relations were expected).

The means and standard deviations of all measures collected in Study 1 are provided in Table 8.

Procedure

The new Aggregated Relational Flexibility measure was developed with items adapted from existing measures of constructs related to flexible thinking, either in general or in a relationship context. The items were formatted as declarative statements involving one's thoughts, feelings, behaviors, communication styles, and ability to take a partner's perspective in different relationship contexts. Participants indicated the extent to which each item was

characteristic of their behavior and attitudes in the past 2 weeks on a 5-point Likert scale (where 1="Not at all characteristic of me", and 5="Extremely characteristic of me") in order to capture the degree of flexibility an individual recently adopted in their relationship (i.e., past 2 weeks) such that the measure would be more sensitive to short term changes. Higher scores indicated a greater degree of flexibility in relationship contexts. For Study 1A, 25 items were shortlisted, of which 11 were reverse scored. All items are provided in Table 1.

In Study 1A, participants completed the new measure, existing measures of general, non-relational, psychological flexibility (Cognitive Flexibility Inventory, Dennis & Vander Wal, 2010), measures of relationship functioning (Couples Satisfaction Index, Funk & Rogge, 2007; Relationship Attribution Measure, Fincham & Bradbury, 1992; Interpersonal Reactivity Index for Couples, Peloquin & Lafontaine, 2010), and a measure of anxiety and depressive symptoms (Depression, Anxiety, and Stress Scales; Lovibond & Lovibond, 1995). Items from the new measure that exhibited strong psychometric properties (see plan for analyses section) were selected and then validated in Study 1B. Study 1B was comprised of 2 parts: in Part 1, participants completed the selected items from the ARF and the same questionnaires administered in Study 1A[1]. In Part 2, participants who had completed Part were invited to provide their email address. These participants were administered the ARF a second time via email, approximately 1-2 weeks after they completed Part 1. In Studies 1C and 1D, the ARF and CSI were administered to couples in two independent samples recruited as part of two separate, larger studies at two other universities.

[1] The PHQ-4, a shorter measure of anxiety and depression symptoms, was administered in Study 1B instead of DASS due to practical session-time constraints.

Plan for Analyses

In Study 1A, an exploratory factor analysis was conducted to examine the internal factor structure of the Aggregated Relational Flexibility measure, using the *psych* package in R (Revelle, 2016). An oblique rotation was used as it was expected that the different ways in which an individual might exhibit flexibility in a relationship would be correlated (e.g., it might be reasonably expected that the ability and willingness to consider different potential behaviors in a relationship might be correlated with ability to adopt a partner's perspective). A scree test (Cattell, 1966) was applied to the plot of successive eigenvalues to determine the number of factors best captured by the ARF. Items with excellent to fair loadings (greater than 0.45) on the extracted factor(s), following the categorization described by Dennis and Vander Wal (2010) and Comrey and Lee (1992), were selected. Additionally, Cronbach's alpha was computed for selected items that loaded on the same factor to ensure that internal consistency within factors was greater than 0.7 (Tavakol, & Dennick, 2011). The subset of items was then used for validity checks and was used as the ARF measure administered in Study 1B.

The selected factors were used to predict established measures of couple's functioning (CSI-4 and RAM) to examine the predictive validity of the new ARF. For convergent validity, the correlations of each ARF factor with general, non-relational, psychological flexibility as measured by the CFI, and with ability to engage in perspective taking as measured by the IRIC were computed. Discriminant validity was established by examining if the ARF was more closely related to general psychological flexibility (CFI) or perspective taking in relationships (IRIC), than general anxiety or depression symptoms (DASS and PHQ-4). Finally, to provide a rigorous test of incremental validity, we examined whether the ARF explained unique variance in couple's satisfaction (CSI-4) after controlling for general psychological flexibility (CFI) or perspective taking in relationships (IRIC).

In Study 1B, a confirmatory factor analysis was conducted on a new dataset (with the same inclusion criteria for participants). Structural Equation Modeling (SEM) was used to examine the fit of data to the model and factor structure of the proposed measure derived from Study 1A. Tests of predictive, convergent, discriminant, and incremental validity, conducted in Study 1A, were repeated with the new dataset. In addition, test-retest reliability was determined by examining correlations between ARF scores obtained in Parts 1 and 2 of Study 1B.

In Study 1C, separate linear mixed-effects models were used to examine the extent to which an individual's ARF predicted their partner's relationship satisfaction (CSI-4), with couple membership as the random intercept to account for nesting of individuals within couples, using separate datasets (Samples 1C-I and 1C-II). The relationship between an individual's ARF and that individual's relationship satisfaction were computed to further validate results obtained in Studies 1A and 1B.

Study 1A Results

Exploratory Factor Analyses

A parallel analysis was conducted on 25 shortlisted ARF items to estimate the number of factors in the data. The analysis revealed two eigenvalues (9.52 and 2.52) greater than 1, and one factor with an eigenvalue of 0.95, indicating that there may potentially be three factors. An exploratory factor analysis with an oblique rotation extracted three factors, which we labeled ARF-Perspective Taking, ARF-Negative Reactions, and ARF-Rigidity. The factors accounted for .33, .13, and .10 of total variance respectively. Items with excellent factor loadings (greater than .70) were retained, and resulted in 10 items in ARF-Perspective Taking (alpha = .94), 4 items in ARF-Negative Reactions (alpha = .89), and 3 items in ARF-Rigidity (alpha = .88). See Table 1 for loadings of all 25-items on each of the 3 extracted factors.

Validity of the ARF measure

Responses to items selected for each factor derived from exploratory factor analysis were added to produce scores on the three subscales. Correlations were computed between each subscale and relationship satisfaction (CSI), tendency to make negative causal attributions in relationships (RAM-Causal), tendency to perceive partners as blameworthy and responsible for negative relationship events (RAM-Responsibility), general psychological flexibility (CFI-10 and CFI-20), perspective taking in relationships (IRIC), and general anxiety and depression symptoms (DASS-A and DASS-D). All correlations are shown in Table 2.

Convergent and Divergent Validity. As expected, ARF-Perspective Taking was strongly positively correlated with general psychological flexibility (.62) and perspective taking in relationships (.82), and moderately positively correlated with relationship satisfaction (.40). Also, as expected, ARF-perspective taking was weakly to moderately negatively correlated with an individual's tendency to make negative causal attributions in relationships (.-.25), tendency to perceive partner as blameworthy and responsible for negative relationship events (-.32), and general measures of anxiety and depression (-.16 for DASS-A; -.35 for DASS-D). The ARF-Negative Reactions scale shows a similar pattern of correlations, with an expected weaker (though still moderate-strong) correlation with perspective taking in relationships (.53). Surprisingly, ARF-Rigidity was only weakly correlated with all measures (all absolute correlations < .27, and mean absolute correlations with all measures was .11). Thus, ARF-Rigidity was dropped from the aggregated measure. The full-scale ARF was computed using the remaining 14 items from ARF-Perspective Taking and ARF-Negative Reactions. All retained items are indicated in Table 1 (bolded items).

Incremental Validity. To examine the incremental validity of ARF, four linear models were computed with relationship satisfaction (CSI) predicted by general cognitive flexibility (CFI-

20) only, perspective taking in relationships (IRIC) only, both CFI-20 and IRIC, and CFI-20, IRIC, and ARF. Model comparisons using ANOVA indicated that while CFI-20 and IRIC each significantly predicted CSI, using both CFI-20 and IRIC resulted in a significant increase in variance explained ($p = .023$). Furthermore, the inclusion of ARF additionally significantly increased the variance explained ($p = .018$) and significantly positively predicted CSI in the full model with CFI-20, IRIC, and ARF as predictors (Standardized B = .30, $p = .02$), indicating that the ARF contains information significant to predicting relationship satisfaction beyond that accounted for by general cognitive flexibility and perspective taking in relationships.

Similarly, ARF significantly incrementally predicted, beyond CFI-20 and IRIC, tendencies to make negative causal attributions (RAM-Causal) and perceive partners as blameworthy or responsible (RAM-Responsibility) when experiencing negative relationship events. Specifically, lower (vs. higher) relational flexibility is associated with more (vs. less) negative causal attributions in relationships (Standardized B = -.28, $p = .038$), and also associated with more (vs. less) perceptions that a partner was blameworthy/responsible for negative relationship events (Standardized B = -.39, $p = .003$). Results of model comparisons are presented in Table 3.

Study 1B Results

Confirmatory Factor Analysis

Confirmatory factor analysis was conducted using Structural Equation Modeling with maximum likelihood parameter estimation. The full hypothesized model, based on results from Study 1A and loadings of items on latent factors, is shown in Figure 1. The model was a good fit to the data collected in Part 1 of Study 1B (Comparative Fit Index = .97; Tucker-Lewis Index, TLI = .96; Root Mean Square Error of Approximation, RMSEA = .054; Chi-square = 172; Restricted Degrees of Freedom = 77). Standardized coefficients of the 10 selected items

loading on the latent ARF-Perspective Taking variable were all above .54 (mean value of coefficients = .85), and standardized coefficients of the 4 selected items loading on the latent ARF-Negative Reactions variable were all above .80 (mean value of coefficients = .89).

Item Reduction

Given that a brief measure is beneficial in terms of reducing participant burden, the number of items in the ARF-Perspective Taking subscale was reduced from 10 to 4. The four items that had the highest coefficients that loaded on the latent variable were retained, and together with the four items that comprised ARF-Negative Reactions subscale, formed the final 8-item full scale ARF. The final set of items retained for the final ARF is indicated in Table 1 (highlighted in gray). A structural equation model with the reduced set of items was tested (see Figure 2), and remains a good fit to the data collected (Comparative Fit Index = .99; Tucker-Lewis Index, TLI = .99; RMSEA = .042; Chi-square = 35; Restricted Degrees of Freedom = 20).

ARF Scale Validity

Construct Validity. Correlations between the ARF subscales and full scale with relationship satisfaction (CSI), general cognitive flexibility (CFI), perspective taking in relationships (IRIC), and anxiety and depression symptoms (PHQ-4) were computed and provided in Table 4. As expected, patterns of correlations with measures of general psychological flexibility, perspective taking in relationships, and anxiety and depression symptoms previously observed in Study 1A were replicated. Specifically, the full-scale ARF as well as the subscales were moderately to strongly positively correlated with general cognitive flexibility (between .40 to .53) and perspective taking in relationships (between .44 to .80). Correlations of ARF (full scale and subscales) with relationship satisfaction were moderately positive (between .23 and .34) and were comparable to correlations of general cognitive

flexibility and perspective taking in relationships with relationship satisfaction (.21 and .29 respectively). As expected, the ARF scales were weakly negatively correlated with symptoms of anxiety and depression (between -.02 and -.15).

Incremental Validity. Similar to the analyses conducted in Study 1A, four models predicting relationship satisfaction (CSI) using CFI only, IRIC only, CFI and IRIC, and CFI, IRIC, and ARF were computed. The four models were then compared to examine the increase in variance explained with the addition of each variable. As expected, the full-scale ARF significantly incrementally predicted relationship satisfaction beyond general cognitive flexibility and perspective taking in relationships. Results showed that the addition of ARF as a predictor to the model significantly increased the variance explained when predicting CSI ($p < .001$). Additionally, ARF was a significant predictor of relationship satisfaction, after accounting for variance explained by CFI and IRIC (Standardized B = .27, $p < .001$). Results from model comparisons are presented in Table 5.

Test-Retest Reliability. Correlations between ARF subscale and full scores were computed for participants who participated in both parts of Study 1B, and are provided in Table 6. As expected, test-retest reliability was good (.8 for the full scale; .8 for ARF-Perspective Taking; .59 for ARF-Negative Reactions).

Study 1C and Study 1D Results

Association between Relational Flexibility and Relationship Satisfaction

As expected, an individual's relationship flexibility significantly positively predicted that individual's relationship satisfaction in both Study 1C (Standardized B = .16, $p = < .001$) and Study 1D (Standardized B = .27, $p < .001$). In other words, individuals who were more (vs. less)

relationally flexible reported higher (vs. lower) satisfaction in their relationship, replicating the results obtained in Study 1A and Study 1B.

Individual's Relational Flexibility and Partner's Relationship Satisfaction

As expected, in Study 1C, an individual's relational flexibility significantly positively predicted their partner's relationship satisfaction (Standardized B = .08, $p < .001$), such that if an individual is more (vs. less) relationally flexible, their partner tended to report higher (vs. lower) relationship satisfaction. However, contrary to expectations, the result was not replicated in Study 1D where no significant relationship was found between an individual's relational flexibility and their partner's relationship satisfaction (Standardized B = .05, $p = .53$). The correlations between an individual's and their partner's relational flexibility and relationship satisfaction for both samples are provided in Table 7. It is important to note that though significant associations were only found in Study 1C, the pattern and magnitude of correlation between an individual's relational flexibility and their partner's relationship satisfaction are similar between Studies 1C and 1D. Furthermore, the standardized B in both models were comparable, but small, indicating that perhaps a significant result was found in Study 1C due to the larger sample size (260 couples in Study 1C vs. 85 couples in Sample 1D).

Summary of Study 1 Results

Taken together, Study 1 established that relational flexibility, as measured by the ARF, was a significant predictor of relationship satisfaction across four different samples. Results indicated that though relational flexibility is closely related to general cognitive flexibility and perspective taking in relationships (as would be expected), relational flexibility provides important, unique information about relationship satisfaction beyond that accounted for by general cognitive flexibility and/or perspective taking. Notably, mixed findings (between Studies

1C and 1D) were obtained regarding the association between an individual's relational flexibility and their partner's relationship satisfaction, though the observed trends were similar across both studies.

Study 2: Piloting Training Materials for Novel CBM-FlexC Paradigm

Methods

Participants

Expert users (N=4) were researchers and clinicians with expertise on couples interventions. End-point users (N=7; 4 male, 3 female) were individuals currently in monogamous romantic relationships (lasting more than 3 months). Due to the qualitative nature of the study design, though power analyses were inappropriate, the sample size was sufficient to iteratively evaluate and improve the clarity and relevance of the CBM-FlexC training materials (following Beard et al., 2011). End-point user participants were paid $50 for participation in the study.

Materials

CBM-Flexible Thinking in Couples (CBM-FlexC). CBM-FlexC is a computerized training program, comprising a number of training scenarios. Each of these training scenarios involves a fictitious couple (the names of the fictitious couples were matched to the gender that each participant in a couple identified with), which remained the same throughout the 6 sessions, but who "experienced" different issues/scenarios in their relationship each session. The participant followed the same two couples, experiencing different issues, throughout the training. See Appendix C for a sample training schedule.

Each training scenario began with a short description of an ambiguous relationship-relevant scenario. Participants were then provided with the same scenario text accompanied by separate sentence stems that provided additional information about the scenario from each partner's perspective. Each sentence stem ended in a word fragment that participants were asked to complete by filling in the missing letter. This was done to ensure active reading and engagement with the material. Given that relational flexibility involves being able to flexibly consider different thoughts, emotions, and behaviors related to relationship-specific events, sentence stems were presented in 3 blocks relating to: (1) thoughts and attributions the couple might have about what might have caused the situation, (2) emotions the couple might feel, and (3) behaviors the couple might engage in. After completing the word fragment, participants were asked to imagine the ways that scenario might unfold given each of the different sentence stems provided. Participants were then asked a Yes/No comprehension question after the imagery part of each sentence stem. After each block of sentence stems, participants were asked to list additional ways the couple might be thinking about the situation, feeling, and positive ways of behaving in the situation.

Problem Domain List. Training materials were designed to fall into one of five broad problem domains in which couples might typically experience conflict, based on the top "problem areas" identified by Miller et al. (2003) and Whisman, Dixon, & Johnson (1997). The domains used in this study were: (1) Communication, (2) Sex/Intimacy, (3) Children/Other family (e.g., in-laws), (4) Finances, and (5) Division of labor (e.g., household chores, childcare). Four different training scenarios were developed per domain.

Attributions, Emotions, and Behaviors in Ambiguous Scenarios. Each scenario comprised a short orienting sentence that described the scenario (e.g., "Chris made dinner for Sophie and himself. They ate in silence"). As described above, each scenario was associated

with 3 blocks of sentence stems, and participants were instructed to imagine the scenario unfolding as vividly as possible given *each* sentence stem. The first block of sentence stems comprised 16 different attributions a partner (8 from the perspective of each person in the couple) might make in the situation (e.g., "Chris thinks "Sophie is still upset about the argument we had yesterday"."). Given the different components of causal attributions described by Fincham and Bradbury (1992), the sentence stems involved positive and negative causal attributions related to the individual him/herself, to the partner, events that occurred in the relationship, and events that occurred outside the relationship. The second block of sentence stems comprised 8 different emotions (2 positive and 2 negative emotions for each partner) that the couple might experience in the situation (e.g., "Chris is hurt by Sophie's silence"). The third block of sentence stems comprised 8 possible behaviors (3 positive and 1 negative behavior for each partner) that might occur in the presented situation (e.g., "Chris shows concern that something seems to be troubling Sophie."). The variety of alternative attributions, emotions, and potential behaviors included in each scenario was designed to stimulate and challenge participants to consider different ways of reacting to a single ambiguous situation, and thus exercise flexible thinking. Further, given that flexibility (not positivity) was the target of training, participants were encouraged to consider positive and negative options related to thinking, feeling, and behaving in these situations.

Comprehension Questions. To ensure that participants adequately attended to the material and understood the resolution to the ambiguity provided by the stem and completed fragment, participants responded ("Yes" or "No") to a comprehension question after each sentence stem. For example, the comprehension question for the sentence stem, "Chris thinks "This meal I made tastes awful"." might be "Did Chris like food he made?"

Images Associated with Ambiguous Scenarios. To facilitate imagery, each scenario was paired with a photograph that was relevant to the context in which the ambiguous situation was taking place. For example, the scenario "Chris made dinner for Sophie and himself. They ate in silence" was paired with an image of plates of food on a dining table.

Assessment Protocol. Participants rated each scenario in terms of relevance to relationship-specific problem domain, degree of experienced ambiguity, and amount of anxiety they would experience in a similar situation. Ratings were made using a 5-point Likert scale (where 1="not relevant at all", and 5="Very relevant"). Participants provided detailed feedback on scenario elements, general comments, and suggestions in qualitative interviews. See Appendix B for the assessment protocol.

Procedure

One-on-one qualitative interviews were conducted with participants. Participants were provided a description of the CBM-FlexC paradigm, training stimuli classified into the appropriate problem domains, a list of ambiguous situations, lists of attributions, emotions, and behaviors related to each ambiguous scenario, and images associated with each ambiguous situation, and the assessment protocol. Participants were asked to review and discuss the assessment protocol and training materials in terms of their clarity, personal relevance, relevance of stimuli to the assigned relationship-specific domain, and interest value. For the purposes of developing new training materials, participants were also asked what other scenarios, attributions, emotions, and behaviors might be relevant to distressed couples. The training materials were iteratively refined and updated based on feedback.

Study 2 Results

Findings from Qualitative Interviews

To guide interviews, a form (see Appendix B) was used that listed topics of interest and rating scales on dimensions such as clarity, relevance, and degree of ambiguity of scenarios. The interviews were conducted using a semi-structured approach to facilitate open-ended discussions that revolved around topics of interests. Participants reviewed scenarios from a minimum of two sets of relationship-specific domains (e.g., communication, intimacy) based on what they identified at the beginning of the study session as being domains that were most relevant to their current romantic relationship. An informal qualitative analysis was completed by the lead author using detailed notes of participant feedback taken during each study session. Key themes in each participant's response, relating to clarity, relevance, and ambiguity of scenarios were identified, and compared between participants to identify commonalities in participants' feedback and generate overall themes.

Clarity and Relevance of Scenarios to Pre-specified Domains of Functioning. Overall, participants reported that the scenarios were clear and relevant to the domains of couples' functioning the scenarios were designed to target (e.g., communication, intimacy, finances). Across all scenarios that were evaluated, the mean participant rating of the relevance of each scenario to the appropriate domain of couples' functioning was 4.4 (on a scale of 1 to 5, where 1 = "not relevant at all" and 5 = "Extremely relevant"; SD = .8). Additionally, participants provided both general and specific feedback to improve the clarity of each scenario, sentence stem, and comprehension question. In particular, participants highlighted the need to reduce the complexity and reading level of materials, because it was generally perceived (prior to revision) that some words and sentences were too complex and/or long. There was also concern from two participants about the frequent use of intense emotions words (e.g., "Sophie *hates* the meal

I made.") and language reflecting extreme all-or-nothing thinking (e.g., "Mike and I can *never* agree on anything."). The sentence stems were thus edited to reduce the use of these intense emotion words and descriptors.

Ambiguity of Scenarios and Expected Anxiety. In general, participants felt that most scenarios adequately described relationship events that were ambiguous and that there were multiple positive and negative attributions one might make in each of the situations. The mean rating for the ambiguity of scenarios was 3.8 (on a scale of 1 to 5, where 1 = "Not ambiguous at all" and 5 = "Extremely ambiguous"; SD = 1.1). In particular, participants felt that including sufficient information about what each couple was doing at the moment in the scenario was important to facilitate imagery and allowed participants to focus on the ambiguity related to the relationship (vs. general ambiguity about what is going on in the situation). For example, changing "Mike comes home with an armful of groceries. Angelica hears the door slam as Mike enters the house." to "Mike comes home with an armful of groceries. *Angelica is working in her home office.* The door slams as Mike enters the house.", allowed participants to focus on the relationship tension elicited in the scenario vs. focusing on what Angelica might be doing at that moment in order to imagine the scenario unfold. Overall, participants rated scenarios as being moderately anxiety provoking if they were to imagine themselves facing a similar situation (M = 3.0, SD = 1.2; where ratings were on a scale of 1 to 5, with 1 = "Not anxious at all" and 5 = "Extremely anxious"). Participants' moderate ratings of scenario ambiguity and expected anxiety were consistent with expectations, given that extreme relationship events (e.g., "Chris heard from a friend that Sophie might be cheating on him.") that tended to be less ambiguous and evoke more intense anxiety were not used.

Relevance of Scenario to Personal Experience. Participants reported that they were able to more easily relate to scenarios that were more similar to their own relationship

experiences. Most participants provided explicit verbal and non-verbal indications when a scenario was very relevant to their own experiences. For example, one participant, upon reading the scenario about a partner leaving dishes in the sink, chuckled and said "That happens all the time!", and proceeded to describe in detail his own experiences, thoughts, and feelings about dividing chores in the household with his partner. Participants generally reported higher levels of anxiety when imagining themselves in situations described in scenarios that were more (vs. less) congruent to their own experiences. Participants also generally found it easier to produce alternative thoughts, feelings, and behaviors that the fictional couple in the scenario might experience if participants had personal experiences with similar situations, compared to if they had never or rarely encountered these situations in their personal life.

Diversity Considerations of Scenarios. The majority of participants commented that the names used for couples should be edited to reflect broader racial/ethnic diversity. For instance, participants suggested changing the names "Jay" and "Betsy," which were used in the original set of scenarios, because these names are frequently associated with White individuals in the United States. A participant suggested the use of shortened names (e.g., Mike vs. Michael) as she perceived shortened names as being more ambiguous in terms of indicating race/ethnicity of an individual. The U.S. Social Security Administration baby names online database (https://www.ssa.gov/oact/babynames/) was used to select names for our revised scenarios based on the popularity of names across states in various years (e.g., popular names in Puerto Rico in the 80's and 90's). Subsequently, the names "Chris", "Mike", "Lucas", "Mateo", "Sophie", "Liz", "Kim", and "Angelica" were selected for our scenarios.

Additionally, several participants commented that a few sentence stems involve situations that could only occur in higher income households. For instance, "Sophie thinks: "Our cleaner will take care of the mess tomorrow."" reflected higher income status that many

participants might not be able to relate to. These sentence stems were revised to improve relatability of scenarios to participants from a wide range of socio-economic backgrounds (e.g., "Sophie thinks: "It will be easy to clean up the mess tomorrow morning."").

Summary of Study 2 Results

Overall, participants reported that scenarios were clear and relevant to the targeted problem domains, and were moderately ambiguous and anxiety provoking when participants imagined themselves being in those situations. Participants reportedly found it easier to relate to scenarios that were more similar to their own experiences. Additionally, participants highlighted the importance of ensuring scenarios reflected a broad range of socio-economic and racial/ethnic diversity.

Study 3: Feasibility and preliminary test of efficacy of the CBM-FlexC paradigm

Methods

Participants

Participants were distressed couples who are currently in monogamous romantic relationships for at least 3 months. Participants were recruited via online advertisements (e.g., Craigslist, Facebook Ads, Reddit Ads, Google Adwords) or recruitment flyers posted in various locations in the Charlottesville/Staunton, Virginia, area. Couples were prescreened using the short 4-item version of the Couples Satisfaction Index (CSI-4; Funk & Rogge, 2007). Those couples with at least one partner scoring at or below the established cutoff that indicates distress in their relationship (score < 13.5) were invited to participate in the study. Couples were excluded if one or both individuals met any of the following exclusion criteria (following Doss et

al., 2016): (1) not in a monogamous romantic relationship, (2) in a monogamous romantic relationship for less than 3 months, (3) in a long-distance relationship, (4) reported more than moderate levels of suicidal ideation (>=7 on the Suicidal Behaviors Questionnaire-Revised; SBQ-R; Osman et al., 2001), (5) reported experiencing intimate-partner violence (>10.5 on the Hurt, Insult, Threaten, and Scream questionnaire; HITS; Sherin et al., 1998), (6) reported making concrete plans to separate or divorce, (7) reported an ongoing affair, (8) reported ongoing couples therapy or plan to seek couples therapy during the study, or (9) had no high speed internet connection.

A total of 494 individuals responded to the recruitment advertisements, 171 of these individuals completed the first online screener, and 53 of the partners to these individuals also completed the partner online screener. Out of these 53 complete pairs of couples who completed the online screener, 38 couples met criteria and were contacted for the phone interview. After completing the phone interview, 23 couples were invited to participate in the study. Of the 23 couples, 3 couples dropped out during baseline (so never started training), 1 couple dropped out after the first training session, and another couple dropped out after the second training session, thus leaving N = 18 couples (18 women; 18 men) who completed the study. The majority of participants who completed the study identified as heterosexual, with one cis-gendered female participant who identified as gender-queer. Participants were between 21 to 62 years old (M = 36.19 years old, SD = 12.4). The majority of participants reported their race/ethnicity as White (32 individuals), while 2 participants reported as Native American/Alaskan Native, and 2 reported as mixed race. All participants identified as non-Hispanic. Couples reported that they were currently "in a relationship" (7 couples), "engaged" (3 couples), or "married" (8 couples). Thirteen couples reported that they had been in a relationship for more than 3 years, 2 couples reported being in their relationship for 1-2 years,

and 3 couples reported that the length of their relationship was between 3 to 12 months. Each couple was paid $100 for participating in and completing the study.

Given that this is a pilot feasibility and efficacy study with a limited time-frame, and the potential for scalability and dissemination of an online intervention to couples that could address issues related to accessibility and stigma, an online sample of couples was used and the current study was conducted via the internet. However, it is important to recognize that with this design, in-person behavioral assessments were not available. For the current project, a multiple baseline design with follow-up was used, given that such a design is ideal for pilot studies with small samples that still want to allow for causal inferences (see Barlow & Hersen, 1984). The multiple baseline design uses each couple's own baseline, established during the baseline period, as their own control, negating the need for a traditional control group. The small sample size used in the present study is typical and appropriate for multiple baseline approaches (e.g., Behrens et al., 1990; Blackwell & Holmes, 2010).

Materials

Demographics information. Participants provided information about their age, gender, race/ethnicity, education, relationship status, length of relationship, and if the relationship is long distance. See Appendix E for items included in the online pre-screening survey.

Suicidal Ideation. The Suicidal Behaviors Questionnaire-Revised (SBQ-R; Osman et al., 2001) is a brief 4-item measure of suicidality. Specifically, lifetime suicidal ideation was measured on a scale of 1 ("Never") to 4 ("I have attempted to kill myself, and really hoped to die."). Frequency of ideation was measured on a 5-point Likert scale (where 1="Never", and 5="Very Often (5 or more times)"). Threat of suicidal behaviors was measured on a scale of 1 ("Never") to 3 ("Yes, more than once, and really wanted to do it"). Likelihood of suicidal behavior was measured on a 6-point Likert scale (where 0="Never", and 6="Very Likely").

Intimate Partner Violence. The Hurt, Insult, Threaten, and Scream questionnaire (HITS; Sherin et al., 1998) is a 4-item screener of frequency of intimate partner violence, using a 5-point Likert scale (1="Never", and 5="Frequently").

General Psychological Flexibility. The Cognitive Flexibility Inventory (CFI; Dennis & Vander Wal, 2010) described in Study 1 was administered.

Relational Flexibility. The 8-item Aggregated Relational Flexibility (ARF) measure developed in Study 1 was administered as the primary outcome measure. It measures the extent to which individuals engage in flexible thinking regarding relationship-specific issues on a 5-point Likert scale (where 1="Not at all characteristic of me", and 5="Extremely characteristic of me").

Aspects of Relational Functioning. The Interpersonal Reactivity Index for Couples (IRIC; Peloquin & Lafontaine, 2010) and the short Couples Satisfaction Index (CSI-4; Funk & Rogge, 2007) described in Study 1 were administered. The Differentiation of Self Index (DSI; Skowron and Friedlander, 1998) is a 43-item questionnaire that measures the extent to which an individual is able to maintain and balance cognitive and affective responses, and the need for intimacy vs. autonomy in close relationships. Items are rated on a 6-point Likert scale (from 1="Not at all true of me", and 6="Very true of me"). The DSI comprises four subscales: (1) Emotional Reactivity reflecting emotional sensitivity to relationship events (e.g., "When someone close to me disappoints me, I withdraw from that person for a time"), (2) I-Position reflecting ability to maintain a clearly defined sense of self even in relationship conflicts (e.g., ""When I am having an argument with someone, I can separate my thoughts about the issue from my feelings about the person"), (3) Emotional Cutoff reflecting feelings of excessive vulnerability in relationships that result in distancing behaviors (e.g., "I tend to distance myself when people get

too close to me"), and (4) Fusion with Others reflecting being overly emotionally involved with others (e.g., "When my partner is away for too long, I feel like I am missing a part of me").

CBM-Flexible Thinking in Couples (CBM-FlexC). The CBM-FlexC training was comprised of the training materials developed in Study 2.

Procedure

Couples responding to the online advertisement first completed a brief online screener and demographics survey. Those eligible, based on the online survey, were contacted by email to schedule a 30-minute phone interview with each partner. Each couple was screened following the criteria outlined in the *Participant* section above. Given that the study was conducted online, establishing initial contact via phone calls allowed verification of participant identity (e.g., verifying couple status), and encouraged accountability and reduced attrition (e.g., Senturia, 1998).

Participants completed at least 3 baseline assessments, each spaced approximately 3 days apart (Barlow et al., 1984). During the baseline assessment phase, participants were emailed a link where they would respond to a set of online questionnaires, including measures of relational and non-relational psychological flexibility, and various aspects of couples functioning. In this study, a stable baseline was established when there was less than 30% variability in the 3 baseline measurements of relational flexibility. Each baseline assessment was administered approximately 3 days apart following the procedure described in Reuland and Teachman (2014). Specifically, we computed the ratio of the maximum difference between any two of the three scores and the highest score (of the three scores). A ratio of less than 0.30 indicated that a stable baseline is established. Once a stable baseline was established, participants were emailed a link to the online CBM-FlexC program.

During the training phase, CBM-FlexC, the computerized training program designed in Study 2, was administered to participants in 6 sessions over 3 weeks. Participants were asked to complete 2 training sessions per week, and each training session took about 15-20 minutes. Participants received email reminders on the day of each session. Participants were reminded before each session that they should complete the training independently from their partners, so that each individual could fully engage in the training without interference from their partner (e.g., a partner who provided their own immediate, biased interpretation of an ambiguous scenario, might constrain or prevent the individual from expressing or working through their own interpretations of the situation). Before the first session, participants were asked to select the 3 areas (out of 5; see problem domain list in Appendix B) in which they experienced the most difficulties in their relationship. In each training session, participants completed 2 training scenarios related to 2 of the domains they had chosen (see Appendix C for sample training schedule). Given that CBM interventions typically involve 1-8 sessions, with multiple training sessions resulting in better outcomes (e.g., Hakamata et al., 2010; Hallion & Ruscio, 2011), we elected to administer 6 sessions (each 15-20 minutes) over 3 weeks for this pilot study. Additionally, participants completed assessments of couple's functioning and relational flexibility (using the ARF developed in Study 1) at baseline, and after sessions 2, 4, and 6 (all online). Finally, approximately one month after session 6, participants completed a follow-up evaluation (online questionnaires and exit phone interview). See Appendix D for measurement plan. In this final evaluation, participants completed the same battery of questionnaires administered pre-intervention (including measures related to relational and general psychological flexibility, and couples functioning). After completing the online questionnaires, participants were contacted via email to schedule a 30-minute phone exit interview with each partner separately. Upon

completion of the phone interview with both partners, a $50 Amazon gift card was emailed to each partner.

Hypotheses and Plan for Analyses

We expected that couples' degree of relational flexibility (as measured by aggregated relational flexibility measure) and relationship satisfaction (as measured by CSI-4) would improve from baseline to post-training and be maintained at 1-month follow-up. It was also expected that, given the link between differentiation of self and relationship satisfaction (Skowron, 2000), DSI indices would also improve from baseline to post-training, and be maintained at follow-up. Additionally, though the intervention was targeted at improving relational flexibility, given the expected link between relational and general cognitive flexibility, it was also expected that general cognitive flexibility might improve post training, though the effects would be smaller than those observed for relational flexibility. We further hypothesized that the extent of improvement of relational psychological flexibility would predict degree of improvement in perceived satisfaction in couples' relationships (by each partner) and mood symptoms. We did not have specific hypotheses about the moderation effect of gender on changes in relational flexibility and relationship satisfaction over time (baseline to training to 1-month follow-up). To test these hypotheses, analyses were performed using linear mixed-effects models and graphical inspection, following Reuland and Teachman (2014).

Graphical inspection, standard for multiple baseline designs, was used to assess changes in key outcome indices (e.g., relational flexibility) over the course of the intervention (Blackwell & Holmes, 2010). Total relational flexibility scores for each participant were plotted over time to infer changes resulting from CBM-FlexC training. Substantial increases in scores on the aggregated relational flexibility measure after the introduction and completion of at least

one CBM-FlexC training session would be indicative of causal links between the proposed intervention and increases in relational flexibility (Kazdin, 2003). To aid systematic visual analyses, visual guides in the form of mean ratings, 2-standard deviation bands, and range bars were plotted for each participant's scores during the baseline and training phases, following recommendations by Manolov et al. (2016). Substantial change in scores during and after the training phase was defined by an increase of more than 2 standard deviations (SD) from mean of scores during the baseline phase based on inspection of the graph. Following Reuland and Teachman (2014), subjective ratings were obtained from 4 independent raters (including the lead author) who were instructed to categorize responses into 4 categories: (1) Treatment Responders: If relational flexibility substantially increased over the course of the intervention, and only after the introduction of the intervention, (2) Improvers: If relational flexibility clearly increased over the course of intervention but with no clear difference between the slopes during baseline and intervention periods, (3) Non-Changers: If relational flexibility did not significantly change between baseline and intervention periods, and (4) Decliners: if relational flexibility decreased substantially following intervention.

As recommended by Manolov et al. (2016), given that the subjectivity of graphical inspection can limit the replicability of results that are based solely on visual analyses, quantitative analysis using mixed-effects models (Lumpkin et al., 2002; Nich & Carroll, 1997) was conducted to model group-level change over time. In addition to allowing changes over time to be modeled, mixed effects models are particularly suited for the current study as these models can account for non-normal and/or non-linear distributions of data that is typical in small n designs (Shadish & Rindskopf, 2007) as well as non-independence between responses provided by individual participants (e.g., individuals in a couple). Mixed-effects models were computed for relational flexibility as the criterion variable, with time (i.e., session number) as a

predictor, and subject nested within couples included as the random intercept to account for variability in mean relational flexibility for each couple and participant. Six planned contrasts were conducted for each model to examine changes in the outcome measures over time. Specifically, the planned contrasts compared mean levels of ARF, CSI, and CFI between the training and baseline phases, between one-month follow-up and baseline, and between one-month follow-up and training. The planned contrasts also examined linear and quadratic trends during the baseline and training phases. A significant positive increase in relational flexibility in the training phase (compared to baseline), and no linear or quadratic trends during the baseline phase, provided evidence for a causal relation between the intervention and changes in relational flexibility. Furthermore, a significant positive change in relational flexibility at one-month follow-up compared to baseline, coupled with at least no (or positive) change in relational flexibility at one-month follow-up (compared to training), provided evidence that relational flexibility was maintained (or further improved) one month after training ended. To examine how changes in relational flexibility affect couples' functioning, the slopes of regression lines of time (number of days) predicting relational flexibility and relationship satisfaction were computed, and reflected change in relational flexibility and relationship satisfaction over the course of the study. Linear models were computed using slope/change in relational flexibility to predict slope/change in relationship satisfaction.

Results

Participant Adherence and Participant Characteristics

All participants completed (at least) three baseline sessions (sessions 1 to 3), six CBM-FlexC sessions (sessions 4 to 9), and two sessions one-month after session 9 (sessions 10 and 11, where session 11 was the phone exit interview). Participants took a mean of 83.8 days (SD

= 20), or about 2.8 months, to complete the study. Participants completed sessions that were at least 3 days apart, and were informed that they should complete each session on time as best as they could. Overall, the mean number of days between sessions (across sessions 1-9) was 6.5 days (SD = 4.0) and the modal number of days between sessions was 5, indicating that most participants managed to complete approximately 2 sessions per week in general, though most took more than 3 days between sessions. The mean number of days between sessions 9 and 10 (one-month follow-up questionnaire) was 31.8 days (SD = 3.3).

All of the 18 couples recruited met the criteria of having at least one partner score below the established cutoff (< 13.5) on the CSI-4 during the online screening session, indicating relationship distress was present in all recruited couples at screening. 15 of these couples continued to meet criteria at baseline 1 (session 1), while the CSI-4 scores of both partners in three couples (couples #C, #N, and #P) were above the cutoff at the first baseline session. That said, at least one partner in each of these 3 couples scored below the cutoff in at least one other session during the baseline period, perhaps pointing to the more volatile nature of relationship satisfaction in couples who are experiencing some form of conflict or stress in their relationship.

Baseline Stability

Stability for each participant over the three baseline assessment sessions was assessed using the formula outlined in the *Procedure* section (i.e., baseline stability was achieved if the ratio of the maximum difference between any two of the three baseline scores and the highest score (of the three scores) was less than 0.30). The variability of each participant's baseline scores on the ARF was computed. All participants who completed the study met the criteria for baseline stability on the ARF and proceeded to the CBM-FlexC training sessions. The mean baseline variability among the 36 participants was .13 (SD = .07).

Quantitative Analyses to Assess Change Over Time

The means and standard deviations of the ARF, CSI, CFI, and four DSI subscales obtained during the baseline, training sessions, and one-month follow-up are presented in Table 9. Prior to analyses, all scores on the measures were scaled using the Proportion of Maximum Scaling (POMS) method as an alternative to standardization (following recommendations by Little, 2013, and Moeller, 2015) in order to preserve differences between individuals and changes in scores within individuals over time. All data from the 18 participating couples were used in the analyses[2]. Outliers were identified as residuals that were greater than 2 standard deviations from the estimated models and were removed.

Changes in Relational Flexibility

As expected, results indicated a significant positive linear change in ARF scores as sessions progressed (from 3 baseline sessions, to the training sessions, and to the one-month follow-up; $B = .087$, $p < .001$), indicating that ARF scores generally increased over the course of the study. See Figure 3 for plot of ARF means across sessions. Planned contrasts indicated that, as expected, ARF scores were significantly higher during the training phase than during the baseline phase ($B = .039$; $p < .001$). There was no significant linear ($B = .008$, $p = .34$) or quadratic change ($B = .004$, $p = .35$) in ARF scores across the three baseline sessions, indicating that ARF stability was generally achieved during baseline. There was also no significant linear ($B = .007$, $p = .45$) or quadratic change ($B = .004$, $p = .41$) in ARF scores across the training sessions, indicating that ARF scores did not change significantly between

[2] Note, one female participant (#P1) informed the first author that her grandmother, to whom she was very attached, had passed the week the participant completed the last session of training. The participant stated that "[her] answers this week are greatly affected by [her grandmother's passing]". A decision was made to still use all her data (rather than omit the score she described as atypical) because it was deemed important to be consistent across participants and there was no exclusion of data due to unusual life events for other participants.

training sessions. There was a significant positive difference in ARF scores at one-month follow up compared to mean scores during training (B = .013, p < .001), such that ARF scores increased in the one-month following training completion. Additionally, all significant trends were observed even after controlling for changes in general cognitive flexibility over time.

Changes in Relationship Satisfaction

As expected, CSI showed a significant positive linear change when looking across all study sessions (B = .053; p < .001). The change in CSI means over time is shown in Figure 4. Planned contrasts were conducted. Results showed no significant linear (B = .01, p = .29) or quadratic (B = .003, p = .64) changes in CSI scores between the 3 baseline sessions, indicating that CSI was stable during the baseline period. Furthermore, as expected, there was a significant positive difference (B = .016, p = .006) in mean CSI scores during training vs. during baseline, indicating that relationship satisfaction improved as a result of training. Results also indicated a significant positive linear trend in CSI scores during training (B = .024, p = .012), such that CSI tended to increase across training sessions. A significant quadratic trend was also observed during training (B = .013, p = .018) such that CSI scores tended to dip toward the end of training sessions though scores did not drop to the levels at the start of training. As expected, there was also a significant positive difference in CSI when comparing scores at one-month follow-up to scores during training (B = .011, p = .012), indicating that CSI was not only maintained but was generally higher one month after training ended. Additionally, all significant trends were observed even after controlling for changes in general cognitive flexibility over time.

Changes in General Cognitive Flexibility

As expected, a significant positive linear change in general cognitive flexibility (CFI) was observed when looking across all study sessions (B = .03, p = .003). The change in CFI means

over time is shown in Figure 5. A significant positive difference was also observed when comparing CFI scores during training vs. baseline (B = .0095, p = .03), such that CFI scores during training were significantly higher than scores during baseline. None of the other planned contrasts yielded significant findings, indicating that CFI scores increased during training, but generally did not change between training sessions and the higher CFI scores were maintained at one-month follow-up.

Changes in Differentiation of Self

As expected, differentiation of self (as measured by DSI subscales) significantly increased due to training (see Figure 6). Note that the DSI was administered only at 3 time points (Baseline: sessions 3, training: session 9, one-month follow up: session 10), which allowed assessment of change from before training to after training, and to one-month follow-up. Significantly lower emotional reactivity (hypersensitivity to others/relationship stimuli) was observed after training compared to baseline (B = -.015, p = .009), and the reduction was maintained from after training to one-month follow-up (B = -.002, p = .73). Similarly, emotional cutoff (tendency to feel threatened by intimacy, and feelings of excessive vulnerability in relationships) and fusion with others (indicating emotional overinvolvement with others) were significantly lower after training compared to baseline (Emotional cutoff: B = -.025, p = .004; Fusion with others: B = -.024, p = .007) with no significant difference between end of training and follow-up at one month (Emotional cutoff: B = -.003, p = .68; Fusion with others: B = -.003, p = .67). Finally, results indicated that scores on "I-position" (sense of self) were significantly higher after training compared to baseline (B = .016, p = .044) and the improvement was maintained at one-month follow-up (B = .001, p = .85).

Exploratory Analyses: Does Sex Matter?

Analyses of changes in ARF and CSI scores across sessions were repeated with sex (coded as a 2 level factor, i.e., male vs. female) as a moderator variable. No significant time x sex interactions were observed in predicting ARF or CSI scores, indicating that changes in relational flexibility or relationship satisfaction over time did not vary as a function of whether a participant was male or female.

Exploratory Analyses: How does an individual's relational flexibility impact their partner's relationship satisfaction?

To examine how individuals' ARF score influenced their partner's CSI score over the course of the study, analyses were conducted with partners' CSI scores as the criterion variable, individual ARF and session number as predictors, and couple and individual participant numbers as the random intercepts to account for differences between couples and nesting of individuals within couples. The 6 planned contrasts used earlier for comparing effects of individual ARF on partner's CSI between study phases were again used in these analyses. A main effect of individual ARF was observed, such that, in general, higher (vs. lower) individual ARF scores were associated with higher (vs. lower) partner CSI scores. Additionally, results indicated a significant session (one-month follow-up vs. training) and individual ARF interaction ($B = .077$, $p = .003$), such that an individual's ARF had a greater impact on their partner's CSI at one-month follow-up compared to during the training phase. Figure 7 shows how an individual's partner's CSI changed over sessions depending on that individual's ARF. Taken together, higher individual relational flexibility is generally associated with higher partner relationship satisfaction, and the positive effect is even stronger at one-month follow-up compared to during training.

To examine how overall change in individuals' ARF or CSI over the course of the study was related to overall change in their partner's CSI, slopes of regression lines of individual ARF and partner CSI across time (number of days from the first baseline session) were computed to provide a measure of how much an individual's ARF or their partner's CSI changed during the study. Change in individuals' ARF was used to predict overall change in their partner's CSI using a mixed-effects model, with couple number as the random intercept to account for differences between couples. Results were not significant ($B = .30$, $p = .08$), indicating that change in an individual's relational flexibility was not significantly associated with the change in relationship satisfaction for their partner. A similar mixed-effects model was computed using change in an individual's ARF to predict their partner's CSI at one-month follow-up. No significant results were found ($B = .19$, $p = .13$).

Taken together, the level of an individual's relational flexibility generally positively impacted their partner's relationship satisfaction, and the effect appeared to be greater particularly one-month after the study (compared to during the training). However, an individual's absolute level of relational flexibility (vs. degree of change) appeared to have a greater impact on their partner's relationship satisfaction, indicating that if an individual's relational flexibility improved as a result of training, but remained at a relatively low level overall, the improvement might not be sufficient to improve their partner's relationship satisfaction.

Graphical Analyses

Relational flexibility, measured by the ARF, was used as the primary outcome measure to examine the efficacy of the CBM-FlexC paradigm. Total ARF scores were plotted for the baseline sessions (sessions 1 to 3), alternate CBM-FlexC training sessions when ARF scores were collected (sessions 5, 7, and 9), and at the one-month follow-up (session 10). Increases in ARF scores after onset of the intervention, but not during the baseline period, would indicate a

potential causal link between the CBM-FlexC intervention and increases in relational flexibility. The first author and three additional raters (who were doctoral level graduate students in Clinical Psychology in the Department of Psychology at the University of Virginia, and who were not involved in the current project) independently reviewed the deidentified graphs for each participant, and categorized each participant into one of the four categories laid out above (i.e., treatment responders, improvers, non-changers, decliners). Visual aids (i.e., means, 2-SD bands, range bars) were included in each plot. A categorization was valid if a clear majority (at least 3 out of 4) of raters agreed on the assignment. Raters met in person to discuss participants for whom there was no clear initial consensus on the categorization in order to determine a final categorization.

When comparing ARF scores during the training phase to baseline scores, 6 participants were rated to be "treatment responders", that is, these participants' ARF scores were significantly higher during the training phase compared to baseline and the improvement occurred only after training. 10 participants were "improvers" in that their ARF scores were higher than baseline scores but there were no clear changes in slope between training and baseline phases. 19 participants were classified as "non-changers", where there was no significant change in scores between training and baseline phases, and 1 participant was rated as a "decliner". When comparing ARF scores at the one-month follow-up to the baseline phase, 19 participants were rated to have higher scores one month after the study, 12 participants were deemed to have similar scores to baseline, and 5 participants were rated as having lower scores than baseline, indicating that slightly more than half of the participants (53%) had higher ARF scores one month after the study compared to baseline. When comparing ARF scores at the one-month follow-up to the training phase, 10 participants had higher scores one month after the study, 22 participants had scores comparable to during the training, and 4 participants

had scores lower than during the training phase, indicating that the majority of participants (89%) maintained (or had higher) scores one month after the study.

When comparing CSI scores between the training and baseline phases, 6 participants were rated to be "treatment responders", 4 were "improvers", 23 were "non-changers", and 3 were "decliners", indicating that the majority of participants were rated to have relationship satisfaction that was similar at the training and baseline phases. When comparing CSI scores at one-month follow-up to baseline, 12 participants were rated to have significantly higher scores one-month after the training compared to baseline, 16 were deemed to have similar scores at one-month follow-up vs. baseline, and 8 were rated to have lower scores at one-month follow-up. When comparing CSI scores at one-month follow-up to training, again the majority of participants (21) were rated to have similar scores between the two time points, 11 were rated to have improved scores, and 4 participants were rated to have lower satisfaction at one-month follow-up vs. training.

Taken together, results seem to indicate that though scores on the ARF and CSI remained largely unchanged from baseline to training phases for a large proportion of participants, a number of participants also saw some form of positive change in their scores, especially one month after training, perhaps indicating that participants might require time to translate and observe changes in flexible thinking gained from training into their lives. See Figure 8 for representative plots of relational flexibility (over time) for participants in each of the four categories.

Qualitative Analysis of Exit Interviews

A script listing important discussion points (Appendix G) was used to guide the exit interviews and encourage open-ended conversations. Detailed notes of the exit interviews were taken, and included direct quotes from conversations with the participants that pertained to their

experiences during the study. An informal qualitative analysis was conducted by the first author based on these notes. Each participant's responses were examined for themes related to their impressions of the study, aspects of the study (e.g., presentation of materials, format of sessions, time commitment) that they found helpful/unhelpful, and perceived changes in relationship functioning during and after the study. Themes identified based on each participant's responses were aggregated across participants, and the common themes that emerged are described below.

Perceived effectiveness of intervention. Overall, the majority (29 out of 36) of participants reported that they found the study useful and noticed improvements in aspects of their relationship during the course of the study. Specifically, these participants reported noticing that the manner in which they communicated with their partner and resolved disagreements improved (e.g., "by asking: help me understand why?", or "we started to puzzle things over together"), and that they more frequently tried taking their partner's perspective (e.g., "being able to stop in the middle [of a conflict] and see things from [partner's] point of view", "if we have an argument or misunderstanding, the program was at the forefront of our minds and we were more likely to picture the situation from each other's perspective", "I would stop and think, maybe it's this or that"). Participants also reported noticing differences in how their partner responded to relationship events (e.g., "[partner] has started doing "I feel this way" more often", "... noticed [my partner] doing more critical thinking"). Additionally, participants reported that when disagreements occurred, one or both partners would often paraphrase the awkward or tense situation in a form that "mimicked" the language scenarios were presented in the study (e.g., "[partner's name] must be thinking: ..." or "[partner's name] feels that ... "), which helped refocus the couple on the objective issue at hand, added some levity to the otherwise tense

situation, and "enabled [the couple] to talk about things [the couple] would not have talked about before".

Interestingly, couples reported that they started reminded each other about the study when they noticed their partner adopting rigid stances in an argument, or felt that they needed space to communicate their thoughts and feelings to their partner. Couples reported that bringing up the study often served as a reminder to "step back", and frequently helped relieve in-the-moment tension that was often followed by adaptive conversations about how and what both partner thought and felt, and productive problem solving. Of 36 participants, 5 reported that they did not notice changes in their relationship, and 2 participants reported that they were uncertain about whether changes occurred. All participants denied that the study had a negative impact on the quality of their relationship.

Relevance of Scenarios. The majority (27 out of 36) of participants felt that many of the scenarios they encountered were relevant and relatable to a certain extent. For example, a participant reported that some of the scenarios they read were "uncanny" and that they could readily imagine or had experienced very similar situations unfolding in their own relationship. Other participants reported ""how did they know [what our issue/difficulty was]?", that "the stories were timely because a lot are scenarios that we experienced", or "some responses were exactly what we would do". However, 9 participants felt that their own life situations are very different from scenarios presented in the study, making it "difficult for [the participant] to see the point of thinking about the situation in different ways" (e.g., participants reported that "the nature of conflicts seemed childish", that they "didn't feel that scenarios were related to anything in my world", or "it felt more like playing a video game rather than being useful"). All participants reported that the relevance and similarity of scenarios to their own experiences were important features to promote visualization and engagement.

Reactions to Training Materials. Participants generally felt that each session was straightforward and that they were able to easily understand their tasks in each section of a session (e.g., filling in word fragments, answering comprehension question, answering the free-response questions). Participants were mixed in their experience with producing alternative, positive thoughts, feelings, and behaviors for each scenario--14 participants felt that the 60-second time limit was too restrictive and induced anxiety and stress in them. Several of these participants reported having difficulties "typing fast enough" and that they "have a hard time coming up with answers".

Participants had differing opinions about the presentation of scenarios from the perspective of fictitious couples. Three participants stated that the scenarios/training would be much more helpful for them if they were guided to think through scenarios "if [they vs. a fictional couple] were the ones experiencing the situation" and reported that "the program was a good first step, but lacks the human touch" and "doing it as a couple might be more helpful". On the other hand, another participant reported preferring the "distance" that was created by thinking about relationship events from the perspective of the fictional couple, stating that the "distance" allowed them to "start to practice" taking perspective without feeling defensive. Similarly, another participant reported that "the third person [perspective] makes it easy to tune down emotions a little".

Related, participants stated that they might have found the training more helpful if some of the interactions they typically engage in with each other regarding the study (e.g., commenting on and discussing scenarios they read in a session, post-session) were integrated into the study design. In particular, a participant suggested the addition of shorter sessions in which the couple would collaboratively brainstorm alternative thoughts, feelings, and behaviors

the fictitious couples might have experienced in situations they previously read about, or discuss ways they would have approached a similar situation in their own relationship.

Length of Sessions and Engagement with Study. Time taken to complete each training session was a major concern for 18 participants. These participants noted that the length of time per session was too long, and coupled with the repetitive nature of the tasks reduced their level of engagement (e.g., they reported that the sessions "were boring and long", "frustrating", "felt repetitive and belaboring the point", "became kind of a drag"), especially in the second half of each session. Additionally, a number of participants reported the time needed for each session discouraged them from completing the sessions during the day (e.g., during lunch time). These participants would often only check their emails at work, but then forget about the study when they got home and had more available time. Ten participants stated that they would have preferred reading about one (vs. two) couple per session, because session time would be reduced and thinking about one couple's situation in greater depth would be helpful. Seven participants reported that 6 training sessions was too many, and two of these participants specifically reported noticing that they "started losing attention and engagement in the 3rd to 4th session".

Completing Sessions on Mobile vs. Desktop Computers. Though participants were encouraged to complete sessions on their personal desktop computers or laptops, 7 participants reported completing the study using only their mobile phones (due to convenience and accessibility). Many participants stated a desire to complete sessions on their mobile phones as that would allow them to access and complete the sessions more easily. All participants who reported using their mobile phones to complete at least part of the study commented on the need to allow word fragment responses to be case-insensitive as phones auto-capitalize letters in the response field. Several of these participants also reported having difficulties completing

their responses for the timed, free-response questions as it was slower to type on mobile phones.

Summary of Study 3 Results

Taken together, for the overall sample, results from the mixed effects models indicated that CBM-FlexC training led to positive changes in relational flexibility, perceived relationship satisfaction, and differentiation of self that are maintained one-month after the training. Furthermore, an individual's absolute level of relational flexibility (vs. extent to which relational flexibility had changed during training) appeared to be more closely tied to their partner's perceived relationship satisfaction. Results from graphical analyses did not indicate positive effects as clearly given the ARF and CSI scores for many participants were deemed not to have reliably changed between the baseline and training phases, though positive trends were observed in a number of cases. Informal qualitative analysis on participant reports at the one-month follow-up indicated concerns about the time required for each session and the repetitiveness of sessions, and also highlighted the importance of "match" between participant relationship concerns and scenarios to promote engagement. Notwithstanding, a majority of participants reported that the training sessions were useful and interesting in terms of stimulating perspective taking and adaptive communication strategies in their relationship.

Discussion

The primary goals of this pilot project were to examine the association between relational flexibility and relationship satisfaction, and to evaluate the feasibility and efficacy of applying a novel CBM intervention, CBM-FlexC, to improve relational flexibility and increase relationship satisfaction. A new aggregated relational flexibility measure (ARF) was developed

and used as one of two primary outcome measures (the second being couple satisfaction index to measure relationship satisfaction) in the CBM-FlexC intervention that comprised 3 baseline sessions, 6 training sessions, and a one-month follow-up. In Study 1, the validity and reliability of the new ARF were established in a 4-part study with different independent populations. As expected, Study 1 indicated that relational flexibility was a significant predictor of relationship satisfaction, even after accounting for the variance explained by general cognitive flexibility and perspective taking. Furthermore, results from quantitative analyses in Study 3 indicated that, as expected, training was associated with improved relational flexibility and relationship satisfaction, and these gains tended to be maintained even one month after the end of the intervention. Exploratory analyses indicated that there were no significant gender differences in responses to training, and an individual's level of relational flexibility typically had a greater connection with their partner's relationship satisfaction at one-month follow-up vs. during the study.

It should be noted that results of graphical analyses in Study 3 were less positive than what was obtained from quantitative analyses. Particularly, graphical analyses indicated that, though positive trends in ARF and CSI were observed following CBM-FlexC training, most participants were not deemed to have experienced reliable improvement in relational flexibility and relationship satisfaction. Additionally, given the small sample size used in Study 3, intervention results should be interpreted with caution as the observed effects were small and variances were large, and even small individual fluctuations in ratings could have potentially non-negligible impact on the aggregated results. Nonetheless, these pilot results are encouraging and demonstrate the preliminary efficacy of the CBM-FlexC paradigm in improving flexible thinking in relationships. It remains important to replicate these findings in larger studies. Furthermore, given that the impact of training might produce changes even one-month after the

study ended, it will also be important that future larger studies include follow up sessions to examine the longer term outcomes of CBM-FlexC.

Composition of Relational Flexibility

Study 1 revealed that relational flexibility, based on the way it was conceptualized (i.e., being able to think and react flexibly in response to relationship events), is comprised of two main subscales. The first subscale, ARF-Perspective Taking, is related to willingness to consider alternative viewpoints and explanations when disagreements occur in relationships. The presence of this factor is unsurprising, given that perspective taking and being able to consider different alternative explanations for experiences is an important component of cognitive flexibility (e.g., Dennis & Vander Wal, 2010) and has been linked to quality of relationships (e.g., Peloquin & Lafontaine, 2010). The second subscale, ARF-Negative Reactions, largely reflects the tendency to react negatively (and can be cognitive, affective, or behavioral reactions) to disagreements or differences in viewpoints that occur in relationships (e.g., "I reacted strongly when my partner contradicted me"). The ARF-Negative Reactions subscale speaks to the importance of not just being able to internally adopt different perspectives when being relationally flexible, but to also be "flexible" in ways that one responds to their partner and behaves interpersonally in-the-moment when disagreements occur, and highlights the *relational and interpersonal* aspect of the construct. Given that emotional reactivity and inability to downregulate negative emotional reactions in-the-moment can have a detrimental impact on interpersonal relationships (e.g., Bloch, Haase, and Levenson, 2014), while effective emotion regulation can promote closeness and support (e.g., Salvatore et al., 2011), it may be important that individuals are able to react in ways that convey to their partner that they were being open to understanding their partner's viewpoints or perspectives as part of resolving their differences.

Additionally, one might imagine that, during a disagreement, being able to regulate one's emotions, react calmly (or less explosively), and exhibit curiosity could provide the time and space for cognitive processes, such as perspective taking, to occur. Alternatively, it could be that being open to and actively considering alternative explanations of situations and being able to easily reframe one's understanding of situations from the perspective of one's partner could potentially reduce the emotional reactivity and intensity of negative emotional and behavioral reactions. Likely, these two processes occur in parallel and influence each other in a feedback loop that leads to more adaptive interpersonal conflict resolution. It would be interesting to examine the relationship between the two subscales more closely in future studies so that we might better understand the ways relational flexibility could potentially impact relationship functioning and satisfaction and highlight productive pathways for interventions.

General vs. Relational Flexibility

Unsurprisingly, our results indicate a consistently moderate positive correlation between general cognitive flexibility and relational flexibility. This positive relationship is expected given that many aspects of general cognitive flexibility, such as being able and open to considering multiple alternative explanations and being able to generate alternative solutions for difficult situations in one's life, are relevant when dealing with more specific issues in relationships. Thus, it is expected that one's general cognitive flexibility will have a significant influence on the quality of one's relationship. Interestingly, our results indicate that there may be an interpersonal component specific to relational flexibility that is not fully captured by general cognitive flexibility. We found that, even after accounting for general cognitive flexibility and perspective taking, that relational flexibility contributes significant additional information to predicting relationship satisfaction. These results suggest that for effective couples functioning, not only must one be able to consider alternatives and perspectives of events that occur in their lives, but that it helps

for this flexibility and willingness to be applied directly in the interpersonal context and be explicitly expressed and communicated, perhaps affectively and behaviorally, to the partner in order to produce additional positive impact on the relationship. For example, one could imagine an individual who is cognitively flexible having greater ability to cope with difficulties occurring in their lives, thus reducing the amount of stress (e.g., stress at work) that might seep into their romantic relationship. Additional benefit to the relationship can potentially be derived if the person is additionally able to consider how their own presentation due to stress within and outside the relationship might impact their partner, and effectively regulate their affective and behavioral expressions (e.g., communicating with their partner about their stress) in their relationship.

Overall Sample versus Individual Change

It is interesting to note the mixed findings derived from the graphical individual-level visual analyses vs. the more robust findings from the mixed-effects models aggregate analyses. It is important to note that graphical analysis can be a particularly stringent method of analysis, given that it relies on changes being large enough to be observable and detectable via visual inspection (Parsonson & Baer, 1978). Thus, subtle changes and trends might easily be missed. Furthermore, though visual cues (e.g., lines indicating means, see Figure 3 for sample plots) were superimposed on individual plots to serve as guides for decision making (based on recommendations in Manolov et al., 2016), these guidelines could have resulted in overly conservative decision making. In particular, requiring the mean of scores in the training phase to be more than two standard deviations from the mean of scores during the baseline phase to qualify as significant change may have been too stringent in this case where standard deviations of scores are large, given that typically only three data points were included in the

baseline phase. As a result, a number of scores were deemed to not have substantially changed in the graphical analysis.

An additional challenge in interpreting the graphical results is that, although we strove to establish and follow a standardized procedure for classifying all individual plots during graphical analyses, there remains a measure of subjectivity. Furthermore, our graphical analyses focused on individual change, and the variation and relationships *within* couples were not considered. Thus, it might be that subtle, aggregated sample-wide trends and effects were missed when examining changes at an individual level, but these were more reliably detected with statistical mixed-effects analyses that accounted for variations within couples. Additionally, individual-level graphical analyses were prone to noise and outliers. For example, one participant reported a significant loss of a loved one right before the last training session of the study, which significantly impacted her scores that week, and she was classified as a "decliner".

That said, one important consideration is that the sample size of our pilot study is small. With only 18 couples, there may be the possibility of spurious significant findings. Therefore, it is important to interpret all findings with caution until replication can occur. Given that the current study was designed to be a pilot feasibility and preliminary efficacy study, the results are encouraging and established some potential benefits of the CBM-FlexC paradigm in improving relational flexibility and even perhaps relationship satisfaction, but clearly, replicating the study with a larger sample will be critical.

Changes in Relational Flexibility over Time and Number of Intervention Sessions

Results indicated that overall, baseline stability was achieved and relational flexibility within individuals was stable over a period of 1.5 to 2 weeks prior to the intervention. Overall, there was a significant increase in relational flexibility as a result of training. Specifically, there was a significant positive linear trend over the course of the study, and the change occurred

only after the baseline phase when CBM-FlexC was introduced. The linear positive trend was mainly driven by mean differences in scores between study phases; that is, when comparing scores during baseline vs. during training and scores during training vs. one-month follow-up. Importantly, although there was a significant change in relational flexibility scores from baseline to training, there were no significant linear trends during the training period. Together, it seemed to indicate that though training did significantly improve relational flexibility and that benefits can be seen even early in training (i.e., by the second CBM training session), that further improvements in relational flexibility might not occur as a function of increasing CBM-FlexC dose. The relative insensitivity of changes in the target construct to dose of CBM interventions might perhaps be reflective of the varied, and sometimes mixed, findings in literature where the reported session number in other CBM interventions is highly variable between studies (typically ranging between one and eight), and in which the frequency of sessions and outcome measures were also varied. It might also be that participants' engagement were reduced after 3 or 4 CBM sessions, thus reducing the benefits that additional sessions might provide.

Furthermore, the improvement in relational flexibility that was observed from training to the one-month follow-up indicates that time to practice the new thinking style might be an important variable to consider in CBM study designs. It might be that time is necessary for individuals to consolidate the gains from CBM training, perhaps through an iterative process of putting skills learned in training to use and experiencing positive real-world outcomes, which then further reinforces use of the strategies and, in turn, further positive changes in relational flexibility and relationship satisfaction. Thus, a “consolidation period” might be an important consideration when designing future CBM studies, for example, to space sessions further apart or to gradually increase the length of time between sessions as the study progress (e.g., 3 days between the first 3 sessions, and 1 week between last 3 sessions). This could provide more

time for participants to consolidate benefits from early training to better utilize later sessions. It is important to note that, though graphical analyses provided more mixed results (vs. quantitative analyses) regarding the immediate improvement of ARF due to training, graphical analyses did indicate that more participants had higher scores at one-month follow-up (vs. baseline) compared to during the training phase (vs. baseline), which may point to the importance of "consolidation" and practice time in CBM studies.

Changes and Stability of Relationship Satisfaction

As expected, relationship satisfaction was higher during the training (compared to baseline) based on quantitative analyses. Although improvement in relational flexibility occurred early on (in the first 2 sessions of CBM-FlexC training), a closer examination of mean levels of relationship satisfaction at various times during the study appeared to suggest that the improvement in relationship satisfaction occurred after the mid-point of training (after session 6) and was maintained at one month follow-up. One reason for the lag between improvement in relational flexibility and improvement in relationship satisfaction might be that couples needed time to integrate and put into practice the new ways of thinking learned from training into their lives and relationships. Further, it might be that improvements and changes in interactions due to increased relational flexibility needed to be sustained over time, and observed by partners, to produce positive effects on relationships.

The observed lag between changes in relational flexibility (the proximal outcome measure that responded quickly, within one or two training sessions, to onset of training) and changes in relationship satisfaction (perhaps being a more distal outcome measure that responded after halfway through training) points to the importance and value of conducting longitudinal studies with potentially more assessments over a longer duration. It also raises intriguing questions about what might be a more "optimal" and flexible approach to assessment,

given the relatively shorter time needed for the proximal measure to respond to training, the relative insensitivity of the proximal measure to number of training sessions, and the time lag necessary for changes in the distal measure to be observed

Additionally, it is important to acknowledge that, the small sample size in Study 3 limited the extent to which more complex statistical analyses could be applied to the data. The current data pointed to potential time-lagged effects that changes in relational flexibility could have on relationship satisfaction. It would be interesting to replicate our findings in a larger, sufficiently powered, follow up study where more sophisticated statistical methods such as dynamical systems modeling might be applied to examine the dynamics of relationship satisfaction and the interaction with relational flexibility over time.

Impact of Relational Flexibility on Partner's Relationship Satisfaction

Results from our samples were mixed in terms of the association between an individual's relational flexibility and their partner's relationship satisfaction. Specifically, an individual's relational flexibility was significantly positively associated with their partner's relationship satisfaction in Study 1C, but not Study 1D, although the magnitude of the correlations between individual relational flexibility and relationship satisfaction were similar in both studies. One reason might be that given the effect size was relatively small, the larger sample size in Study 1C (which was approximately 3 times larger than Study 1D) could have produced the significant result due to higher power. However, interestingly, a positive main effect of an individual's relational flexibility on their partner's satisfaction was observed in the small Study 3 sample.

A possible explanation might be that the difference between studies was driven by cultural factors inherent in the samples. Particularly, Study 1C (and likewise, the samples used in Studies 1A, 1B, and 3) was comprised of mostly White participants, and the racial/ethnicity demographics composition was more similar to what is generally observed in online studies

(Shapiro et al., 2013). On the contrary, two-thirds of Study 1D participants identified as Hispanic. Thus, it is plausible that cultural factors could have influenced the way in which one thinks about and evaluates their relationships, or how couples engage in conflict resolution in their relationships (e.g., Wheeler, Updegraff, & Thayer, 2010). For instance, there could be a cultural difference tied to greater emphasis on familial values and community cohesiveness in Latino/Hispanic cultures vs. emphasizing emotional connection just between the partners. Specifically, the importance that Latino/Hispanic families place on family closeness, obligations, and solidarity (e.g., Cauce & Domenech-Rodriguez, 2002) might mean that these couples use different personal and cultural guidelines in evaluating their satisfaction in a romantic relationship. Instead of evaluating relationship satisfaction based mainly on their emotional connection with their partner (which we expect would be facilitated by perspective taking), these couples might emphasize the extent to which themselves and/or their partner can fulfill culturally-specified roles in the relationship. This possible difference in the determinants of relationship satisfaction could potentially reduce the impact of relational flexibility on relationship satisfaction, though this suggestion is clearly speculative.

Interestingly, our results indicated that the level of relational flexibility of one partner had a greater impact on the other partner's satisfaction at one-month follow-up compared to during and before training, such that at one month follow-up, higher (vs. lower) relational flexibility was more strongly associated with higher partner satisfaction. This might be because, through participating in the study, participants became more attuned to aspects of their relationship functioning that they might previously not have been aware of, and thus became more sensitive to the presence or absence of certain relationally adaptive behaviors their partners engage in, that in turn impacted their ratings of relationship satisfaction. Surprisingly though, *changes* in relationship flexibility did not significantly predict changes in or the level of relationship

satisfaction in their partners. Perhaps the absolute level of relational flexibility might need to be above a certain threshold in order to impact relationship satisfaction in partners, regardless of whether relational flexibility improved due to training.

Possible Mechanisms Linking the Intervention to Relationship Outcomes

The study was designed to be completed separately by each individual in a couple, and participant exit interviews indicated that all couples adhered to the instructions. However, a majority of couples reported that the study sessions stimulated interesting conversations between themselves in the intervening time between sessions, as they would often discuss and reflect on stories they had previously read in session, especially if the stories were similar to their own experiences. Couples reported that the reflection process provided the opportunity for them to hear and be exposed to the perspectives of their partner, and improved their communication. Additionally, a number of couples reported using the training sessions as prompts to de-escalate real conflicts they encountered. Individuals reported that they would paraphrase their own real-life conflict in-the-moment in a way reminiscent of how scenarios were presented in the study. These participants reported that doing so often provided the cue for both partners to "step back" from how they had reacted to the situation and "think about the situation from [their partner's] perspective", and introduced a moment of levity in what might have been an emotionally tense situation.

Some or all of the above could have contributed in different ways across individuals and couples to produce the observed improvements in relational flexibility and relationship satisfaction. It is intriguing to speculate about the downstream effects of changes in the hypothesized mechanism--promoting thinking about different alternative explanations or experiences in a situation--or other mechanisms that may have been operating. It could be that the exercise of thinking through alternatives during the sessions made it easier for individuals to

accept suggestions and "bids" by their partner to "step back" and think about a situation from different perspectives (e.g., when their partner uses "[individual's name] feels annoyed by the situation" as a cue to engage in perspective-taking exercise). Or, it could be that the sessions provided individuals with simple language to express their thoughts and emotions non-defensively, or to express their understanding of their partner's thoughts and emotions. Or, it could be that collaboratively discussing their thoughts about particular relationship events they read provided each person a greater understanding of the ways their partner might react in different situations that then resulted in changes in their relationship. Unfortunately, in the current study, it is difficult for us to tease apart the individual contributions of these different pathways to improvements in the outcomes. Nonetheless, these results highlight exciting ways in which the CBM-FlexC paradigm might be tweaked and improved to maximize its benefit. For example, individual sessions might be interspersed with sessions that both couples complete together, and in those sessions, one could ask couples to imagine themselves as the characters in a scenario and discuss ways they might resolve that difficult relationship situation.

Furthermore, it might also be that the observed improvements in relational flexibility and relationship satisfaction occurred independently of the actual training sessions (though use of the multiple baseline design, which establishes a stable baseline, make this less likely). Several participants stated that they found the questionnaires helpful in their relationship as the questions prompted them to introspect on their own reactions in situations. Additionally, merely participating in a study together might have given the couples a shared experience that also communicated to each partner that they were both committed to improving their relationship. That said, stability in the measures during the baseline period and the multiple baseline design indicated that merely participating in the study and/or completing only questionnaires (without

the training sessions) was not likely to be sufficient to produce the observed significant changes in relational flexibility and relationship satisfaction.

Technological Considerations

Though participants were instructed to complete the study on their personal computers or laptops as much as possible, a number of participants used their mobile phones in at least a few of the sessions. Even those participants who completed the study solely on their computers reported that they would have preferred to access the study via their mobile phones. This feedback points to the importance of considering mobile phone-based dissemination of treatment, as mobile phones are becoming more easily accessible and more ubiquitous than computers. Implementing and designing CBM interventions for mobile phones means that these interventions would be more easily available for individuals to access when necessary (e.g., to facilitate just-in-time adaptive interventions; Nahum-Shani, Hekler, & Spruijt-Metz, 2015), but also means that CBM sessions must necessarily be quick and in an easily consumable format. For example, the vast majority of participants felt that each CBM-FlexC session covered too much material and was too long (averaging 30 minutes), and would prefer sessions that were at most 10 minutes. Feedback from participants also highlighted the need for mobile-centered design of interventions vs. simply making web-based applications available on mobiles. Important considerations when adapting an online intervention to mobile phones include length of texts visible on the smaller mobile phone screen, amount of text input needed given that it is generally harder for participants to provide a lot of text input with the limited screen size and on-screen keyboard, placement of buttons, and length of sessions. It would be interesting to consider shortening CBM-FlexC sessions, or even have initial sessions comparable to the current session length, but have shorter subsequent sessions that are spread out over a longer time to act as "boosters" or reminders over time.

Limitations and Conclusions

There are several methodological limitations to acknowledge, and it is important that current findings be interpreted in light of these limitations. First, given that this is the first pilot study of the CBM-FlexC paradigm to improve relational flexibility and relationship satisfaction, we elected not to have a control group. Though this is typical of studies using the multiple baseline approach with small samples, the inclusion of a comparison group (e.g., waitlist control and/or a treatment-as-usual group) would provide valuable information about the relative effectiveness and change resulting from the proposed intervention. Follow-up studies incorporating these comparison groups will be valuable in further refining the CBM-FlexC paradigm and for identifying particular features of the paradigm that might be particularly important in producing positive change (e.g., examining the degree to which the online-only CBM-FlexC intervention is similarly effective to in-person treatment as usual).

Second, the study used a small sample of distressed couples who self-selected into the online study. These couples were for the most part motivated to complete the study and improve their relationships, and their choice to complete an online research study may make them different in some ways than other distressed couples. Thus, it will be important to replicate any findings in larger samples that are more representative of the population at large. Additionally, it is plausible that the mixed findings regarding the relationship between individual relational flexibility and partner's relationship satisfaction might be due to cultural and demographic differences, which raises intriguing questions about couples functioning in different ethnic and racial groups and whether there would be differential response to the intervention based on different aspects of participants' identities. Therefore, it will be important in future studies to have sufficient power to examine the impact of cultural variables on relational flexibility, relationship satisfaction, and CBM-FlexC treatment response.

Finally, the potentially diverse sample of distressed couples used in the study may be facing very different relationship challenges, which may or may not make certain training stimuli more relevant than others. For instance, a retiring couple in their 60's would face very different challenges than a newlywed couple in their early 20's. While the proposed intervention attempts to address this concern by providing training stimuli in five broad domains typically implicated in couples functioning/satisfaction and allowing couples to select the domains that are most important to them, it would be interesting to examine, in a future larger sample, the degree to which personalization and match between training stimuli and particular concerns of the couple impact treatment effectiveness or treatment program retention-rates, and explore ways to adapt training stimuli to improve treatment responsiveness of a couple (for example, using Item Response Theory; Fincham & Rogge, 2010).

Despite these limitations, the study is a novel application of CBM, an internet-based paradigm to address a key cognitive mechanism--relational flexibility--in distressed couples. The study is the first to validate and directly measure relational flexibility in couples through the use of the new Aggregated Relational Flexibility (ARF) measure, and is the first attempt to our knowledge at adapting the CBM paradigm to increase relational flexibility in distressed couples. The study has significant implications for addressing relationship distress in a population known to exhibit low-treatment seeking behavior and high treatment dropout rates. These results are a crucial first step in developing an evidence-based intervention technology that has the potential to overcome stigma often associated with couples therapy, and that is easily accessible and non-threatening, thus acting as a promising resource for distressed couples.

References

Barlow, D. H., and Hersen, M. (1984). Single case experimental design, *Pergamon Press,* New York.

Barlow, D. H., Hayes, S. C., & Nelson-Gray, R. O. (1984). The scientist practitioner: Research and accountability in clinical and educational settings (Vol. 128). *Pergamon Press*.

Baucom, D. H., Epstein, N. B., Kirby, J. S., & LaTaillade, J. J. (2002). Cognitive-Behavioral couple therapy. *Handbook of Cognitive-Behavioral Therapies*, 411, 92.

Baucom, D. H., Epstein, N., Sayers, S. L., & Sher, T. G. (1989). The role of cognitions in marital relationships: Definitional, methodological, and conceptual Issues. *Journal of Consulting and Clinical Psychology*, 57(1), 31.

Beard, C. (2011). Cognitive bias modification for anxiety: Current evidence and future directions. *Expert Review of Neurotherapeutics,* 11(2), 299-311.

Beard, C., Weisberg, R. B., & Primack, J. (2012). Socially anxious primary care patients' attitudes toward Cognitive Bias Modification (CBM): A qualitative study. *Behavioural and Cognitive Psychotherapy*, 40(5), 618.

Behrens, B. C., Sanders, M. R., & Halford, W. K. (1990). Behavioral marital therapy: An evaluation of treatment effects across high and low risk settings. *Behavior Therapy*, 21(4), 423-433.

Blackwell, S. E., & Holmes, E. A. (2010). Modifying interpretation and imagination in clinical depression: A single case series using cognitive bias modification. *Applied Cognitive Psychology*, 24(3), 338-350.

Bond, F. W., Hayes, S. C., Baer, R. A., Carpenter, K. M., Guenole, N., Orcutt, H. K., Waltz, T. & Zettle, R. D. (2011). Preliminary psychometric properties of the Acceptance and Action Questionnaire–II: A revised measure of psychological inflexibility and experiential avoidance. *Behavior Therapy*, 42(4), 676-688.

Braithwaite, S. R., & Fincham, F. D. (2007). ePREP: Computer based prevention of relationship dysfunction, depression and anxiety. *Journal of Social and Clinical Psychology*, *26*(5), 609-622.

Burns, D. D., & Sayers, S. L. (1992). Development and validation of a brief relationship satisfaction scale. *Unpublished manuscript*.

Cascardi, M., & Vivian, D. (1995). Context for specific episodes of marital violence: Gender and severity of violence differences. *Journal of Family Violence*, 10(3), 265-293.

Cauce, A. M., & Domenech-Rodriguez, M. (2002). Latino families: Myths and realities. *Latino children and families in the United States: Current research and future directions*, 3-25.

Christensen, A., Atkins, D. C., Baucom, B., & Yi, J. (2010). Marital status and satisfaction five years following a randomized clinical trial comparing traditional versus integrative behavioral couple therapy. *Journal of Consulting and Clinical Psychology*, 78(2), 225.

Christensen, H., Griffiths, K. M., & Farrer, L. (2009). Adherence in internet interventions for anxiety and depression: systematic review. *Journal of Medical Internet Research*, 11(2), e13.

Christensen, A., & Heavey, C. L. (1999). Interventions for couples. *Annual Review of Psychology*, 50(1), 165-190.

Cuijpers, P., Van Straten, A., & Andersson, G. (2008). Internet-administered cognitive behavior therapy for health problems: a systematic review. *Journal of behavioral medicine*, *31*(2), 169-177.

Doss, B. D., Benson, L. A., Georgia, E. J., & Christensen, A. (2013). Translation of integrative behavioral couple therapy to a web-based intervention. *Family Process*, 52(1), 139-153.

Doss, B. D., Rhoades, G. K., Stanley, S. M., & Markman, H. J. (2009). Marital Therapy, Retreats, and Books: The Who, What, When, and Why of Relationship Help-Seeking. *Journal of Marital and Family Therapy*, 35(1), 18-29.

Fadardi, J. S., & Cox, W. M. (2009). Reversing the sequence: reducing alcohol consumption by overcoming alcohol attentional bias. *Drug and Alcohol Dependence*, 101(3), 137-145.

Fincham, F. D., & Bradbury, T. N. (1992). Assessing attributions in marriage: the relationship attribution measure. *Journal of Personality and Social Psychology*, 62(3), 457.

Finn, C., Mitte, K., & Neyer, F. J. (2013). The relationship-specific interpretation bias mediates the link between neuroticism and satisfaction in couples. *European Journal of Personality*, 27(2), 200-212.

Funk, J. L., & Rogge, R. D. (2007). Testing the ruler with item response theory: increasing precision of measurement for relationship satisfaction with the Couples Satisfaction Index. *Journal of Family Psychology*, 21(4), 572.

Georgia, E. J., & Doss, B. D. (2013). Web-based couple interventions: Do they have a future?. *Journal of couple & relationship therapy*, *12*(2), 168-185.

Gottman, J. M., Coan, J., Carrere, S., & Swanson, C. (1998). Predicting marital happiness and stability from newlywed interactions. *Journal of Marriage and the Family*, 5-22.

Greenberg, L. S., & Pascual-Leone, A. (2006). Emotion in psychotherapy: A practice-friendly research review. *Journal of Clinical Psychology*, 62(5), 611-630.

Hakamata, Y., Lissek, S., Bar-Haim, Y., Britton, J. C., Fox, N. A., Leibenluft, E., Ernst, M. & Pine, D. S. (2010). Attention bias modification treatment: a meta-analysis toward the establishment of novel treatment for anxiety. *Biological Psychiatry*, 68(11), 982-990.

Halford, W. K., Keefer, E., & Osgarby, S. M. (2002). “How has the week been for you two?” Relationship satisfaction and hindsight memory biases in couples' reports of relationship events. *Cognitive Therapy and Research*, 26(6), 759-773.

Hallion, L. S., & Ruscio, A. M. (2011). A meta-analysis of the effect of cognitive bias modification on anxiety and depression. *Psychological Bulletin*, 137(6), 940.

Henderson, A. D., Sayger, T. V., & Horne, A. M. (2003). Mothers and sons: A look at the relationship between child behavior problems, marital satisfaction, maternal depression, and family cohesion. *The Family Journal*, 11(1), 33-41.

Holmes, E. A., Mathews, A., Mackintosh, B., & Dalgleish, T. (2008). The causal effect of mental imagery on emotion assessed using picture-word cues. *Emotion*, 8(3), 395.

Huston, T. L., Niehuis, S., & Smith, S. E. (2001). The early marital roots of conjugal distress and divorce. *Current Directions in Psychological Science*, 10(4), 116-119.

Kashdan, T. B., & Rottenberg, J. (2010). Psychological flexibility as a fundamental aspect of health. *Clinical Psychology Review*, 30(7), 865-878.

Kelly, A., Fincham, F. D., & Beach, S. R. (2003). Emerging perspectives on couple communication. *Handbook of Communication and Social Interaction Skills*, 723-752.

Kreider, R.M., & Fields, J.M. (2002). Number, timing, and duration of marriages and divorces: 1996 (Current Population Reports P70-80). *Washington, DC: U.S. Census Bureau*.

Kurdek, L. A. (2004). Are gay and lesbian cohabiting couples really different from heterosexual married couples?. *Journal of Marriage and Family*, 66(4), 880-900.

Lumpkin, P. W., Silverman, W. K., Weems, C. F., Markham, M. R., & Kurtines, W. M. (2003). Treating a heterogeneous set of anxiety disorders in youths with group cognitive behavioral therapy: A partially nonconcurrent multiple-baseline evaluation. *Behavior Therapy*, 33(1), 163-177.

MacLeod, C., & Mathews, A. (2012). Cognitive bias modification approaches to anxiety. *Annual Review of Clinical Psychology*, 8, 189-217.

Magnuson, S., & Shaw, H. E. (2003). Adaptations of the multifaceted genogram in counseling, training, and supervision. *The Family Journal*, 11(1), 45-54.

Marigold, D. C., Holmes, J. G., & Ross, M. (2010). Fostering relationship resilience: An intervention for low self-esteem individuals. *Journal of Experimental Social Psychology*, *46*(4), 624-630.

Masi, M. V., Miller, R. B., & Olson, M. M. (2003). Differences in dropout rates among individual, couple, and family therapy clients. *Contemporary Family Therapy*, 25(1), 63-75.

McGoldrick, M., Gerson, R., & Shellenberger, S. (1999). Genograms. Assessment and intervention. *New York: Norton*.

Miller, R. B., 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(5), 395-407.

Nahum-Shani, I., Hekler, E. B., & Spruijt-Metz, D. (2015). Building health behavior models to guide the development of just-in-time adaptive interventions: A pragmatic framework. *Health Psychology*, *34*(S), 1209.

Nich, C., & Carroll, K. (1997). Now you see it, now you don't: a comparison of traditional versus random-effects regression models in the analysis of longitudinal follow-up data from a clinical trial. *Journal of Consulting and Clinical Psychology*, 65(2), 252.

O'Farrell, T. J., Hooley, J., Fals-Stewart, W., & Cutter, H. S. (1998). Expressed emotion and relapse in alcoholic patients. *Journal of Consulting and Clinical Psychology*, 66(5), 744.

Olson, D. H. (2000). Circumplex model of marital and family sytems. *Journal of Family Therapy*, 22(2), 144-167.

Reuland, M. M., & Teachman, B. A. (2014). Interpretation bias modification for youth and their parents: A novel treatment for early adolescent social anxiety. *Journal of Anxiety Disorders*, 28(8), 851-864.

Revelle, W. (2016) psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, *https://CRAN.R-project.org/package=psych Version = 1.6.12*.

Rothermund, K., Voss, A., & Wentura, D. (2008). Counter-regulation in affective attentional biases: a basic mechanism that warrants flexibility in emotion and motivation. *Emotion*, *8*(1), 34-46.

Rouquette, A., & Falissard, B. (2011). Sample size requirements for the internal validation of psychiatric scales. *International Journal of Methods in Psychiatric Research*, 20(4), 235-249.

Sillars, A., Roberts, L. J., Leonard, K. E., & Dun, T. (2000). Cognition during marital conflict: The relationship of thought and talk. *Journal of Social and Personal relationships*, 17(4-5), 479-502.

Skowron, E. A. (2000). The role of differentiation of self in marital adjustment. *Journal of Counseling Psychology*, 47(2), 229.

Snyder, D. K., Castellani, A. M., & Whisman, M. A. (2006). Current status and future directions in couple therapy. *Annual Review of Psychology*, 57, 317-344.

Steinman, S. A., & Teachman, B. A. (2010). Modifying interpretations among individuals high in anxiety sensitivity. *Journal of Anxiety Disorders*, 24(1), 71-78.

Weiss, R. L., & Perry, B. A. (1979). Assessment and treatment of marital dysfunction. *Eugene, OR: Oregon Marital Studies Program.*

Wells, T. T., & Beevers, C. G. (2010). Biased attention and dysphoria: Manipulating selective attention reduces subsequent depressive symptoms. *Cognition & Emotion*, 24(4), 719-728.

Wheeler, L. A., Updegraff, K. A., & Thayer, S. M. (2010). Conflict Resolution in Mexican-origin couples: Culture, gender, and marital quality. *Journal of Marriage and Family*, *72*(4), 991-1005.

Whisman, M. A. (2001). Marital adjustment and outcome following treatments for depression. *Journal of Consulting and Clinical Psychology*, 69(1), 125.

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(3), 361.

Whisman, M. A., & Uebelacker, L. A. (2006). Impairment and distress associated with= relationship discord in a national sample of married or cohabiting adults. *Journal of Family Psychology*, 20(3), 369.

Tables

Table 1. *Factor loadings of 25 shortlisted Aggregated Relational Flexibility (ARF) items on the factors Perspective Taking (PT), Negative Reactions (NR), and Rigidity (RI).*

		Factors		
	ARF Items	**ARF-PT**	**ARF-NR**	**ARF-RI**
19	**I tried to understand my partner better by imagining how things look from his/her perspective.**	**0.87**	-0.02	-0.01
22	**When I was upset at my partner, I tried to "put myself in his/her shoes" for a while.**	**0.86**	-0.01	-0.02
18	**I thought about an issue from my partner's point of view before I made a decision.**	**0.84**	0.00	-0.01
23	**Before criticizing my partner, I tried to imagine how I would feel if I were in his/her place.**	**0.81**	0.03	-0.03
1	**When my partner and I disagreed on something, I looked at the issue from many different angles.**	**0.81**	-0.01	-0.01
21	**I believed that there are two sides to every issue in my relationship and tried to look at them both.**	**0.80**	0.09	-0.14
7	**I was aware that there could be many reasons when my partner was not responsive to me.**	**0.76**	-0.10	0.15
2	**I considered multiple options before making a decision relating to my relationship/partner.**	**0.75**	0.05	0.03
4	**When I had a conflict with my partner, I realized that I could react in many possible ways.**	**0.74**	-0.03	0.12
8	**I sought additional information before coming to a conclusion about why my partner behaved a certain way.**	**0.71**	-0.16	0.10
	I made sure I had all the facts before reacting to relationship-related issues.	0.67	0.08	-0.08
	When discussing important relationship issues with my partner, I could communicate an idea to him/her in different ways.	0.61	0.14	-0.01
	I was always able to understand my partner's thoughts/feelings without checking with him/her.	0.49	0.05	-0.10
	I was able to take advice from my partner.	0.48	0.22	-0.08
	I adjusted easily to changes in my relationship.	0.42	0.34	-0.15
15	**I did not like hearing opinions from my partner that go against my way of thinking. (-)**	0.02	**0.84**	0.08
14	**I reacted negatively to criticism from my partner. (-)**	-0.06	**0.81**	-0.07
16	**I reacted strongly when my partner contradicted me. (-)**	0.02	**0.81**	0.02
17	**When my partner and I had different opinions, I found it difficult to compromise. (-)**	0.08	**0.71**	0.13
	When I was sure I was right about something, I did not need to listen to my partner's arguments. (-)	0.32	0.40	0.19
6	**There was often only one solution worth considering for handling a problem in my relationship. (-)**	0.05	0.01	**0.86**

5	**When I am feeling indecisive about an issue in my relationship, I felt that there was only one way right way to respond. (-)**	-0.04	-0.01	**0.83**
3	**There was only one way to resolve a difficult relationship issue I was confronted with. (-)**	0.00	0.11	**0.80**
	I felt that a relationship has to be perfect in order to work out right. (-)	-0.10	0.15	0.39
	I preferred to keep things unchanged in my relationship. (-)	0.07	0.13	0.18

Note: Bolded, numbered items are items with factor loadings greater than .70 (in Study 1A) that were retained for confirmatory factor analysis (Study 1B). Items highlighted in gray are those that were retained after item-reduction in Study 1B, and form the final 8-item ARF.

Table 2. *Correlation matrix for exploratory factor analysis (Study 1A)*

	Measure	1	2	3	4	5	6	7	8	9	10	11	12
1	ARF-PT	-											
2	AR-NR	0.45	-										
3	ARF-RI	0.14	0.36	-									
4	ARF	0.95	0.70	0.23	-								
5	CFI-20	0.62	0.44	0.20	0.65	-							
6	CFI-10	0.55	0.41	0.23	0.58	0.97	-						
7	IRIC	0.82	0.53	0.18	0.84	0.58	0.53	-					
8	CSI	0.40	0.41	0.02	0.46	0.36	0.34	0.42	-				
9	RAM-C	-0.25	-0.37	-0.04	-0.33	-0.13	-0.16	-0.32	-0.52	-			
10	RAM-R	-0.32	-0.38	-0.12	-0.39	-0.20	-0.22	-0.35	-0.50	0.80	-		
11	DASS-A	-0.16	-0.40	-0.27	-0.26	-0.42	-0.39	-0.22	-0.36	0.19	0.21	-	
12	DASS-D	-0.35	-0.47	-0.05	-0.44	-0.52	-0.50	-0.36	-0.55	0.36	0.30	0.61	-

Note: ARF-PT = ARF-Perspective Taking; ARF-NR = ARF-Negative Reactions; ARF-RI = ARF-Rigidity; ARF = ARF full scale; CFI = Cognitive Flexibility Index; IRIC = Interpersonal Reactivity Index for Couples; CSI = Couples Satisfaction Index; RAM-C = Relationship Attribution Measure-Causal; RAM-R = Relationship Attribution Measure-Responsibility; DASS-A = Depression, Anxiety, and Stress Scales-Anxiety; DASS-D = Depression, Anxiety, and Stress Scales-Depression. Given that ARF-RI had weak correlations with all other measures and was dropped, ARF was computed using the sum of responses for the ARF-PT and ARF-NR subscales only.

Table 3. *Results from model comparisons examining incremental validity of ARF in predicting couples satisfaction and negative relationship attributions (Study 1A)*

Incremental validity of ARF predicting CSI
Model 1: CSI ~ CFI
Model 2: CSI ~ IRIC
Model 3: CSI ~ CFI + IRIC
Model 4: CSI ~ CFI + IRIC + ARF

Residual Degrees of Freedom	Residual Sum of Squares	Degrees of Freedom	Sum of Squares	F	p	
194	167.69					
194	159.22	0	8.4640			
193	155.08	1	4.1449	5.2838	0.02260	*
192	150.61	1	4.4645	5.6913	0.01802	*

Incremental validity of ARF predicting RAM-C
Model 1: RAM-C ~ CFI
Model 2: RAM-C ~ IRIC
Model 3: RAM-C ~ CFI + IRIC
Model 4: RAM-C ~ CFI + IRIC + ARF

Residual Degrees of Freedom	Residual Sum of Squares	Degrees of Freedom	Sum of Squares	F	p	
194	191.47					
194	175.09	0	16.3877			
193	174.33	1	0.7543	0.8497	0.35779	
192	170.45	1	3.8863	4.3777	0.03773	*

Incremental validity of ARF predicting RAM-R
Model 1: RAM-R ~ CFI
Model 2: RAM-R ~ IRIC
Model 3: RAM-R ~ CFI + IRIC
Model 4: RAM-R ~ CFI + IRIC + ARF

Residual Degrees of Freedom	Residual Sum of Squares	Degrees of Freedom	Sum of Squares	F	p	
194	187.29					
194	173.39	0	13.8961			
193	173.39	1	0.0016	0.0018	0.96584	
192	165.73	1	7.6622	8.8768	0.00326	*

Note: ARF = Aggregated Relational Flexibility measure; CFI = Cognitive Flexibility Index (20 item version); IRIC = Interpersonal Reactivity Index for Couples; CSI = Couples Satisfaction Index; RAM-C = Relationship Attribution Measure-Causal; RAM-R = Relationship Attribution Measure-Responsibility.

Table 4. *Correlation matrix for confirmatory factor analysis (Study 1B)*

	Measure	1	2	3	4	5	6	7	8	9
1	ARF (14 items)	1.00	0.94	0.95	0.91	0.61	0.32	0.52	0.82	-0.02
2	ARF (8 items)	0.94	1.00	0.80	0.84	0.79	0.34	0.53	0.77	-0.07
3	ARF-PT (10 items)	0.95	0.80	1.00	0.95	0.32	0.25	0.45	0.79	0.04
4	ARF-PT (4 items)	0.91	0.84	0.95	1.00	0.33	0.23	0.47	0.80	0.03
5	ARF-NR (4 items)	0.61	0.79	0.32	0.33	1.00	0.33	0.40	0.44	-0.15
6	CSI	0.32	0.34	0.25	0.23	0.33	1.00	0.21	0.29	-0.22
7	CFI	0.52	0.53	0.45	0.47	0.40	0.21	1.00	0.44	-0.25
8	IRIC	0.82	0.77	0.79	0.80	0.44	0.29	0.44	1.00	-0.01
9	PHQ4	-0.02	-0.07	0.04	0.03	-0.15	-0.22	-0.25	-0.01	1.00

Note: ARF = ARF full scale; ARF-PT = ARF-Perspective Taking; ARF-NR = ARF-Negative Reactions; CSI = Couples Satisfaction Index; CFI = Cognitive Flexibility Index; IRIC = Interpersonal Reactivity Index for Couples; PHQ4 = Patient Health Questionnaire-4 (measuring anxiety and depression)

Table 5. *Results from model comparisons examining incremental validity of ARF (8-item version) in predicting couples satisfaction (Study 1B)*

Incremental validity of ARF predicting CSI
Model 1: CSI ~ CFI
Model 2: CSI ~ IRIC
Model 3: CSI ~ CFI + IRIC
Model 4: CSI ~ CFI + IRIC + ARF

Residual Degrees of Freedom	Residual Sum of Squares	Degrees of Freedom	Sum of Squares	F	p	
406	388.41					
406	372.13	0	16.2806			
405	368.52	1	3.6054	4.0657	0.04442	*
404	358.26	1	10.2596	11.5694	0.00074	***

Note: ARF = Aggregated Relational Flexibility measure; CFI = Cognitive Flexibility Index; IRIC = Interpersonal Reactivity Index for Couples; CSI = Couples Satisfaction Index; RAM-C = Relationship Attribution Measure-Causal; RAM-R = Relationship Attribution Measure-Responsibility.

Table 6. *Correlation matrix of ARF full scale and subscale scores at times 1 and 2 (Study 1B)*

	Time 2				
Time 1	ARF (14 items)	ARF (8 items)	ARF-PT (10 items)	ARF-PT (4 items)	ARF-NR (4 items)
ARF (14 items)	**0.80**	0.75	0.79	0.78	0.49
ARF (8 items)	0.76	**0.75**	0.72	0.72	0.56
ARF-PT (10 items)	0.76	0.68	**0.80**	0.79	0.37
ARF-PT (4 items)	0.73	0.66	0.77	**0.76**	0.36
ARF-NR (4 items)	0.52	0.59	0.40	0.41	**0.59**

Note: ARF = ARF full scale; ARF-PT = ARF-Perspective Taking; ARF-NR = ARF-Negative Reactions

Table 7. *Correlations between an individual's relational flexibility and their partner's relational flexibility and relationship satisfaction*

	Self CSI	Partner ARF	Partner CSI
Study 1C: Self ARF	0.27	0.05	0.22
Study 1D: Self ARF	0.27	0.44	0.18

Note: ARF = Aggregated Relational Flexibility measure, full scale; CSI = Couples Satisfaction Index

Table 8. *Means and Standard deviations on measures in Study 1*

	Study 1A		Study 1B		Study 1C-I		Study 1C-II	
Measure	M	SD	M	SD	M	SD	M	SD
ARF (14 items)	50.42	11.04	49.19	9.49	-	-	-	-
ARF (8 items)	29.29	6.53	28.47	5.69	24.05	5.4	29.5	5.31
ARF-PT (10 items)	34.87	8.82	34.05	7.92	-	-	-	-
ARF-PT (4 items)	13.74	3.93	13.33	3.68	-	-	-	-
ARF-NR (4 items)	15.55	3.77	15.15	3.27	-	-	-	-
CSI	14.52	5.03	15.59	4.44	9.46	4.7	17.33	3.14
CFI-20	108.23	16.65	-	-	-	-	-	-
CFI-10	54.72	8.91	56.29	7.89	-	-	-	-
IRIC	16.2	5.3	14.57	4.84	-	-	-	-
RAM-C	44.02	10.93	-	-	-	-	-	-
RAM-R	41.78	11.97	-	-	-	-	-	-
DASS-A	2.12	2.8	-	-	-	-	-	-
DASS-D	3.52	4.59	-	-	-	-	-	-
PHQ4	-	-	2.53	2.58	-	-	-	-

Note: ARF = Aggregated Relational Flexibility full scale; ARF-PT = ARF-Perspective Taking; ARF-NR = ARF-Negative Reactions; CSI = Couples Satisfaction Index; CFI = Cognitive Flexibility Index; IRIC = Interpersonal Reactivity Index for Couples; RAM-C = Relationship Attribution Measure-Causal; RAM-R = Relationship Attribution Measure-Responsibility; DASS-A = Depression, Anxiety, and Stress Scales-Anxiety; DASS-D = Depression, Anxiety, and Stress Scales-Depression; PHQ4 = Patient Health Questionnaire-4.

Table 9. *Means and standard deviations of measures in Study 3*

	Screening (Session 0)		Baseline (Sessions 1-3)		Training (Session 5)		Training (Session 7)		Training (Session 9)		One month followup (Session 10)	
Measures	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
ARF	22.56	4.91	26.20	4.85	27.39	5.48	28.28	5.71	27.89	5.49	28.61	5.24
CSI	11.89	3.29	13.17	3.30	12.81	3.90	14.22	3.49	13.33	3.81	14.33	4.31
CFI	-	-	51.31	8.60	50.81	10.78	53.78	8.83	51.36	8.73	52.56	8.11
DSI-ER	-	-	40.5	6.83	-	-	-	-	38.50	6.62	38.47	6.32
DSI-IP	-	-	44.89	7.12	-	-	-	-	46.25	7.91	46.53	7.38
DSI-EC	-	-	38.14	7.07	-	-	-	-	36.78	7.53	35.61	8.33
DSI-FO	-	-	36.25	5.82	-	-	-	-	34.67	5.53	34.25	5.04

Note: ARF = Aggregated Relational Flexibility full scale; CSI = Couples Satisfaction Index; CFI = Cognitive Flexibility Index; DSI = Differentiation of Self Index; DSI-ER = DSI-Emotional Reactivity; DSI-IP = DSI-"I"-Position; DSI-EC = DSI-Emotional Cutoff; DSI-FO = DSI-Focus on Others. DSI was only measured at Session 3 of baseline phase.

Figures

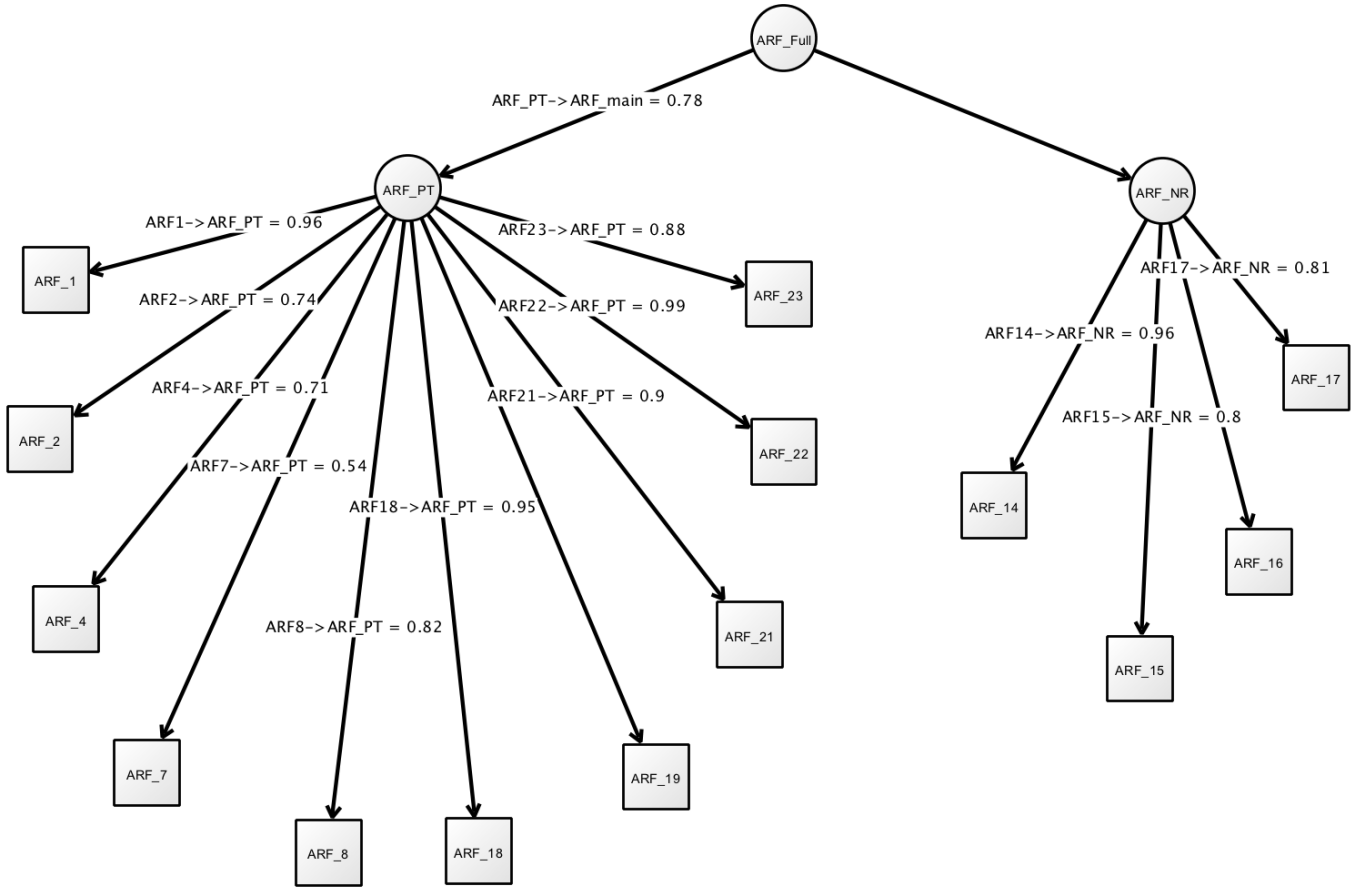


Figure 1. Structural Equation Modeling (SEM) model confirming the factor structure of the ARF (14 items). Loadings of items on latent factors are shown. 10 items loaded on ARF-Perspective Taking (PT), and 4 items loaded on ARF-Negative Reactions (NR).

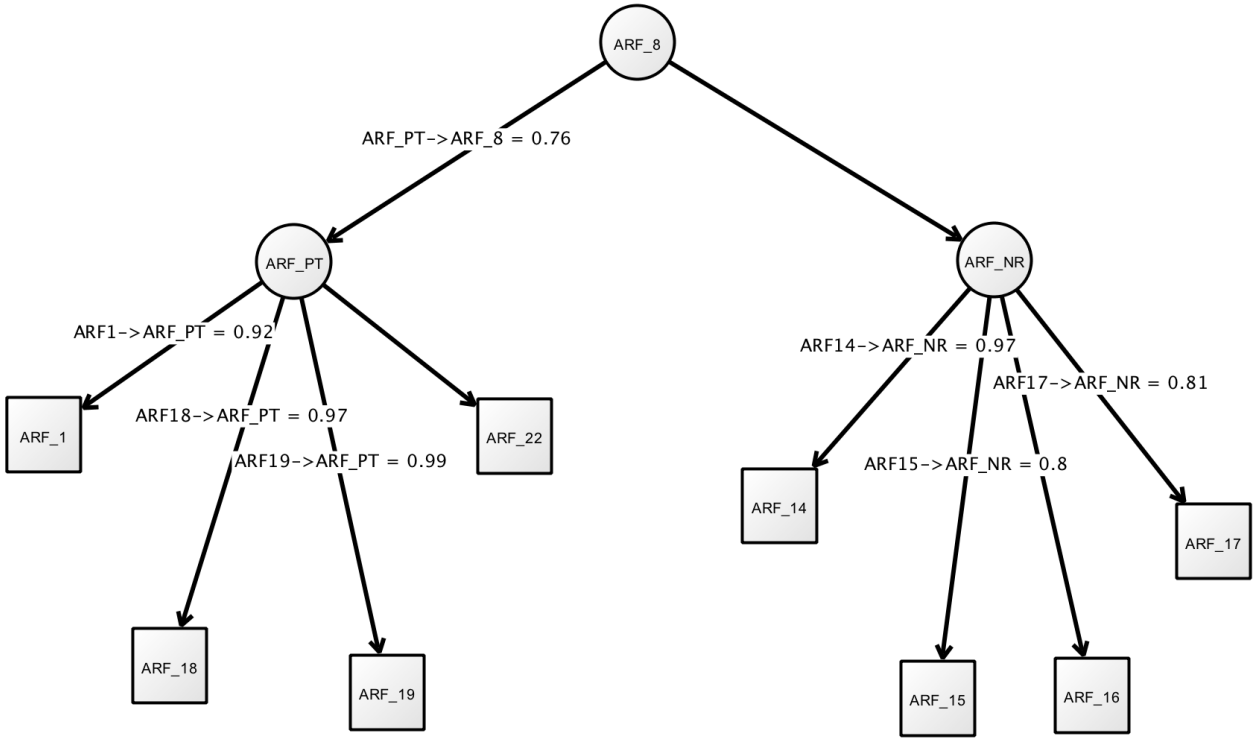


Figure 2. Structural Equation Modeling (SEM) model of the 8-item ARF.
The loadings of 4 items on ARF-Perspective Taking (PT), and 4 items on ARF-Negative Reactions (NR) are shown.

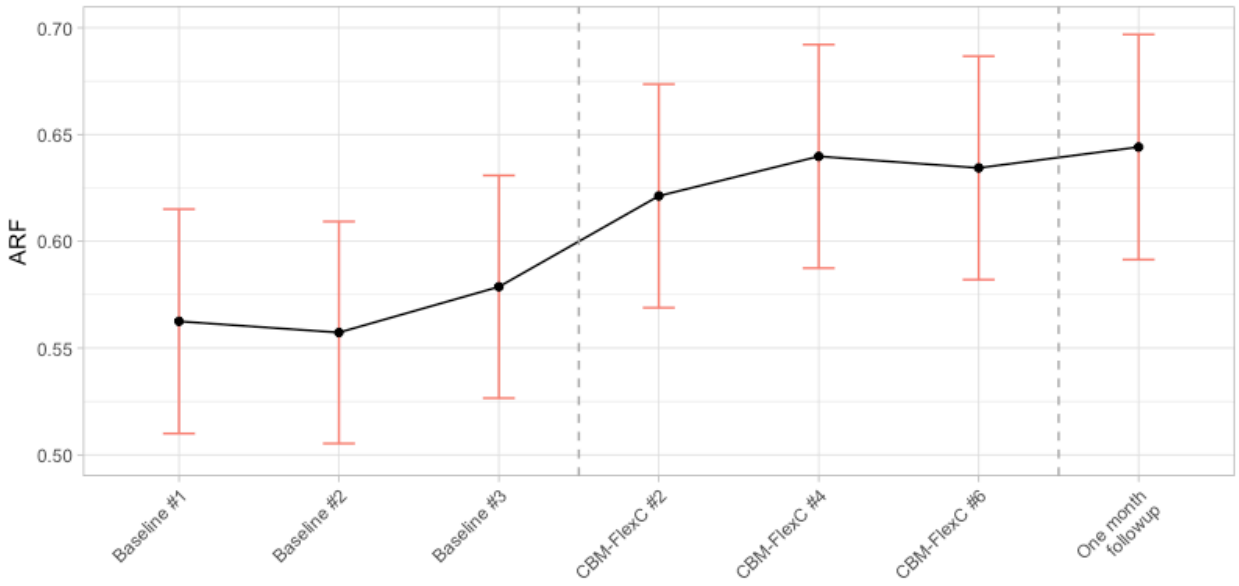


Figure 3. Change in mean levels of relational flexibility over sessions.
ARF = Aggregated Relational Flexibility measure.

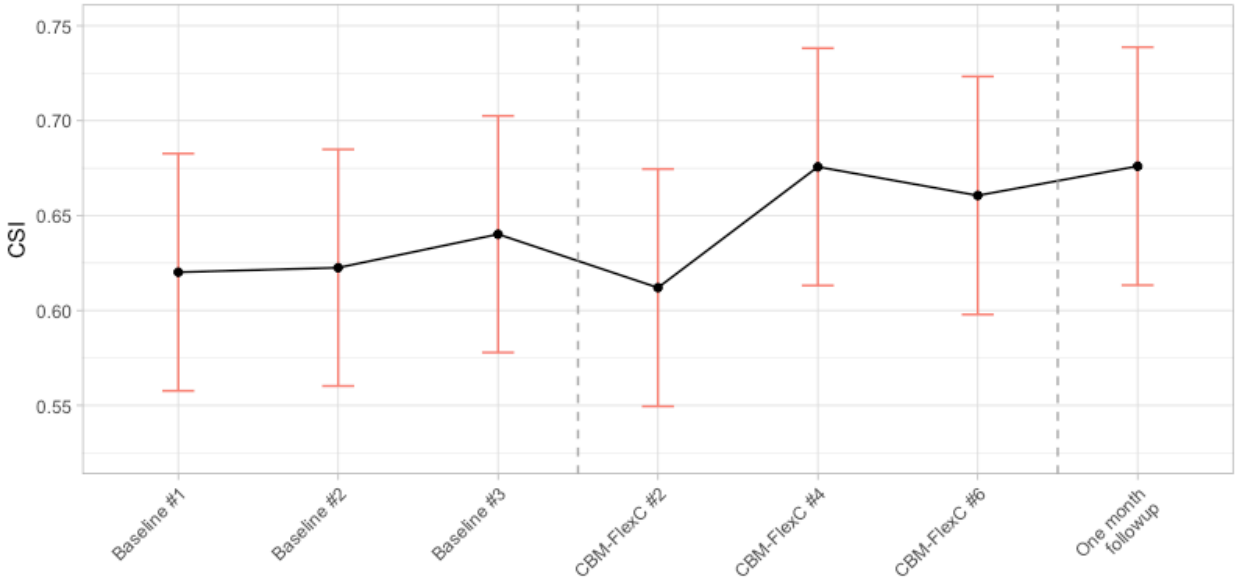


Figure 4. Change in mean levels of relationship satisfaction over sessions.
CSI = Couples Satisfaction Index.

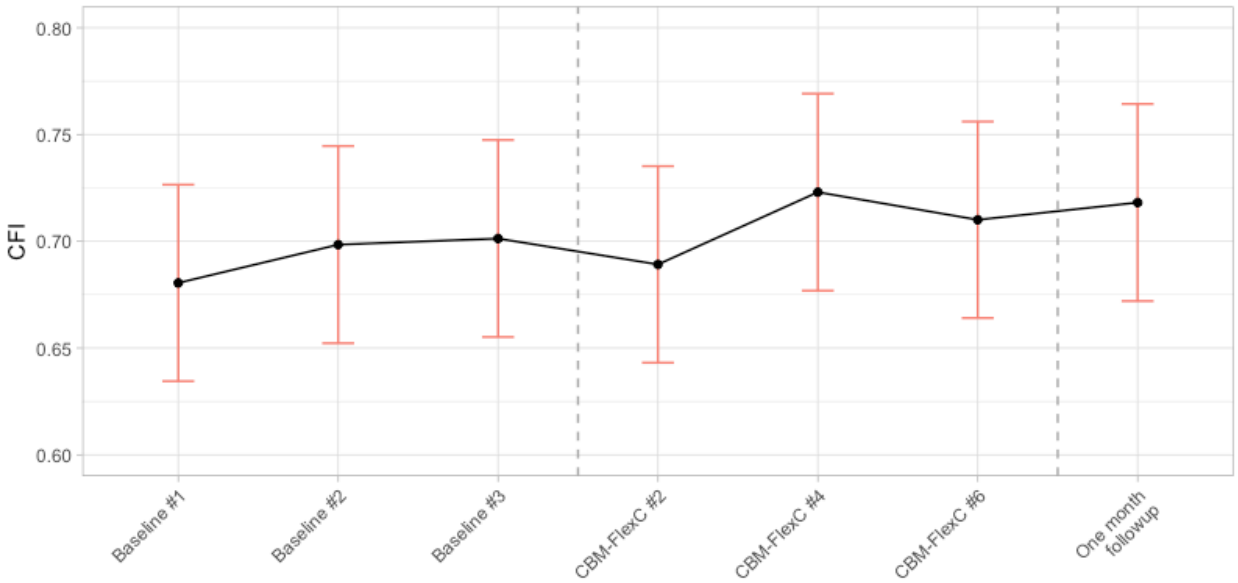


Figure 5. Change in mean levels of general cognitive flexibility over sessions. CFI = Cognitive Flexibility Inventory.

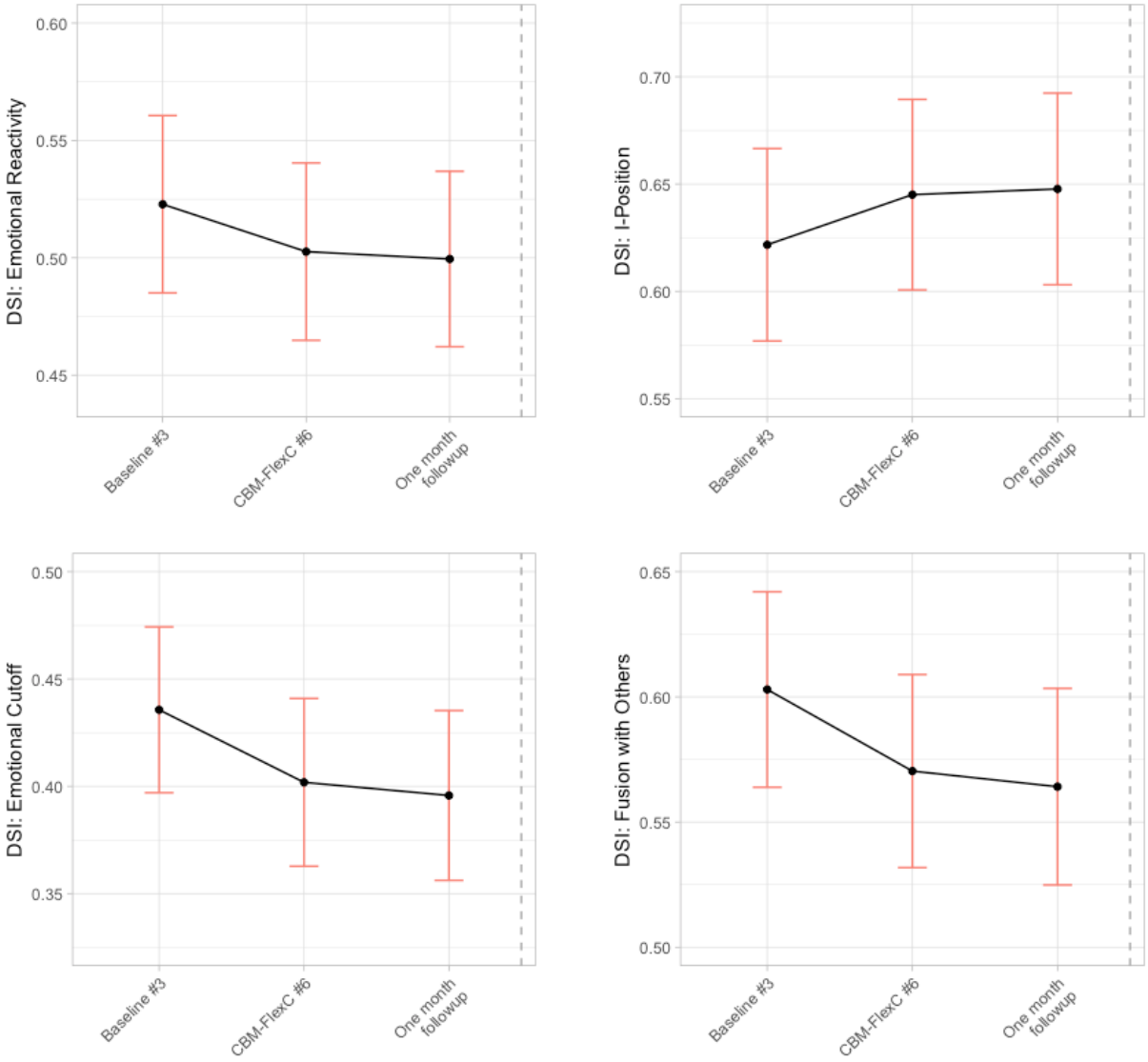


Figure 6. Changes in mean levels of the 4 differentiation of self subscales over sessions. DSI = Differentiation of Self Index.

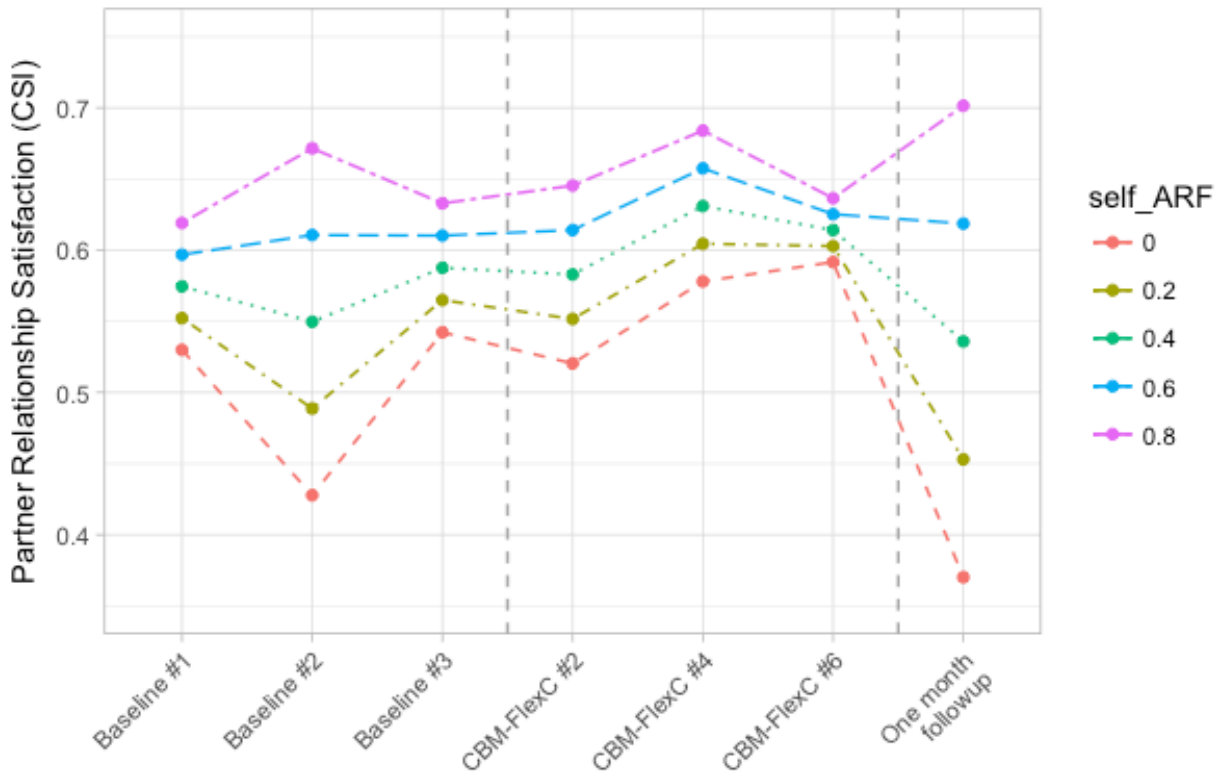
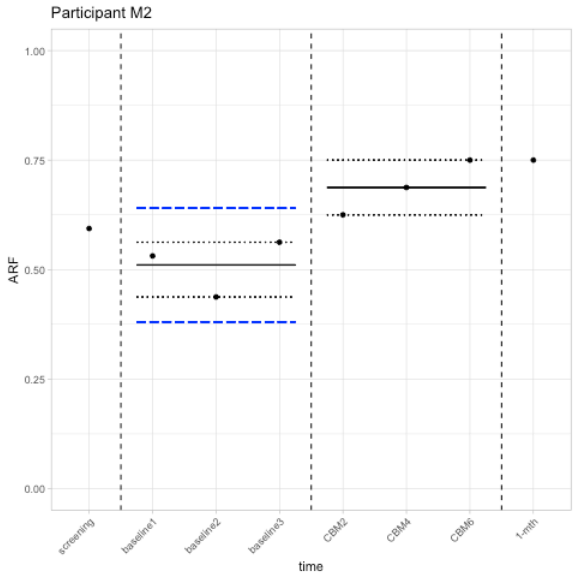
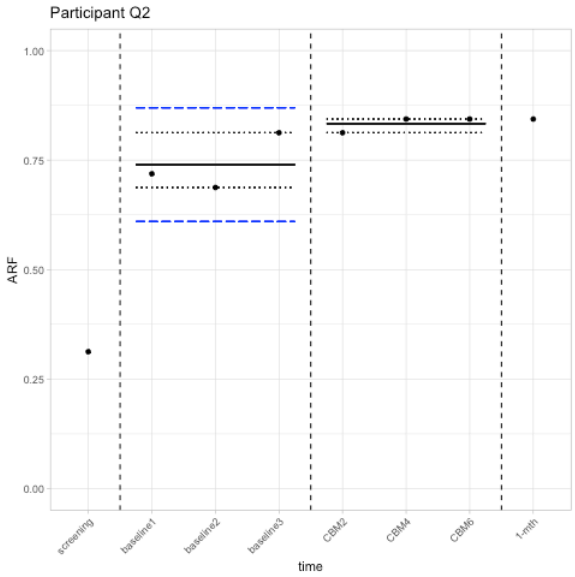


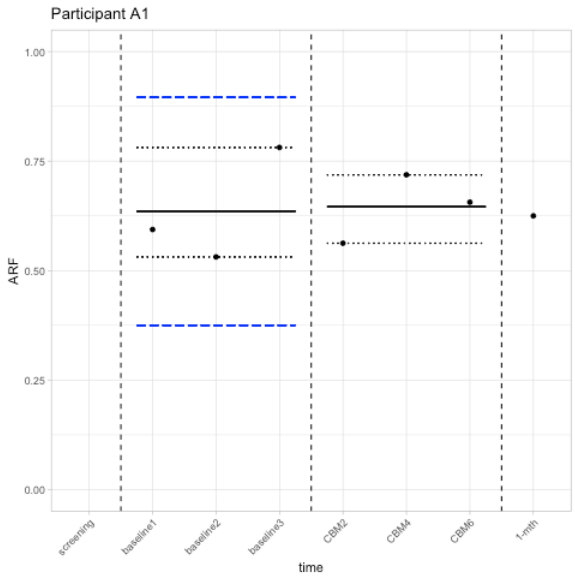
Figure 7. Changes in partner's relationship satisfaction as a function of session and an individuals' relational flexibility.
ARF = Aggregated Relational Flexibility measure; CSI = Couple Satisfaction Index.



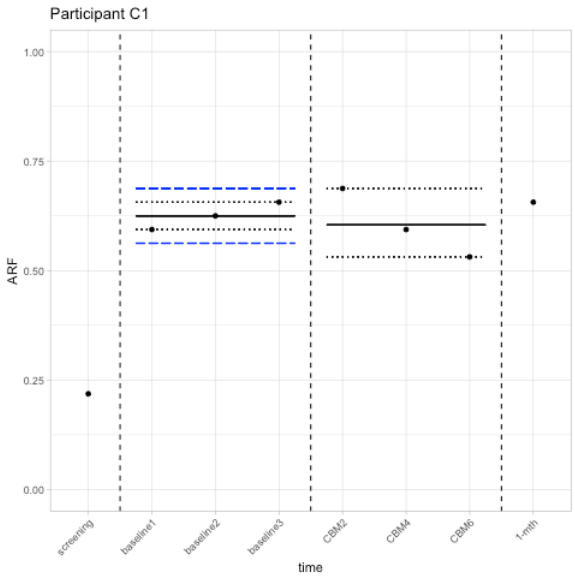
(a) Treatment Responder. Increase in slope of response during training vs. baseline.



(b) Improver. Increase in scores, but no clear increase in slope during training vs. baseline.



(c) Non-changers. No clear change in scores between training vs. baseline.



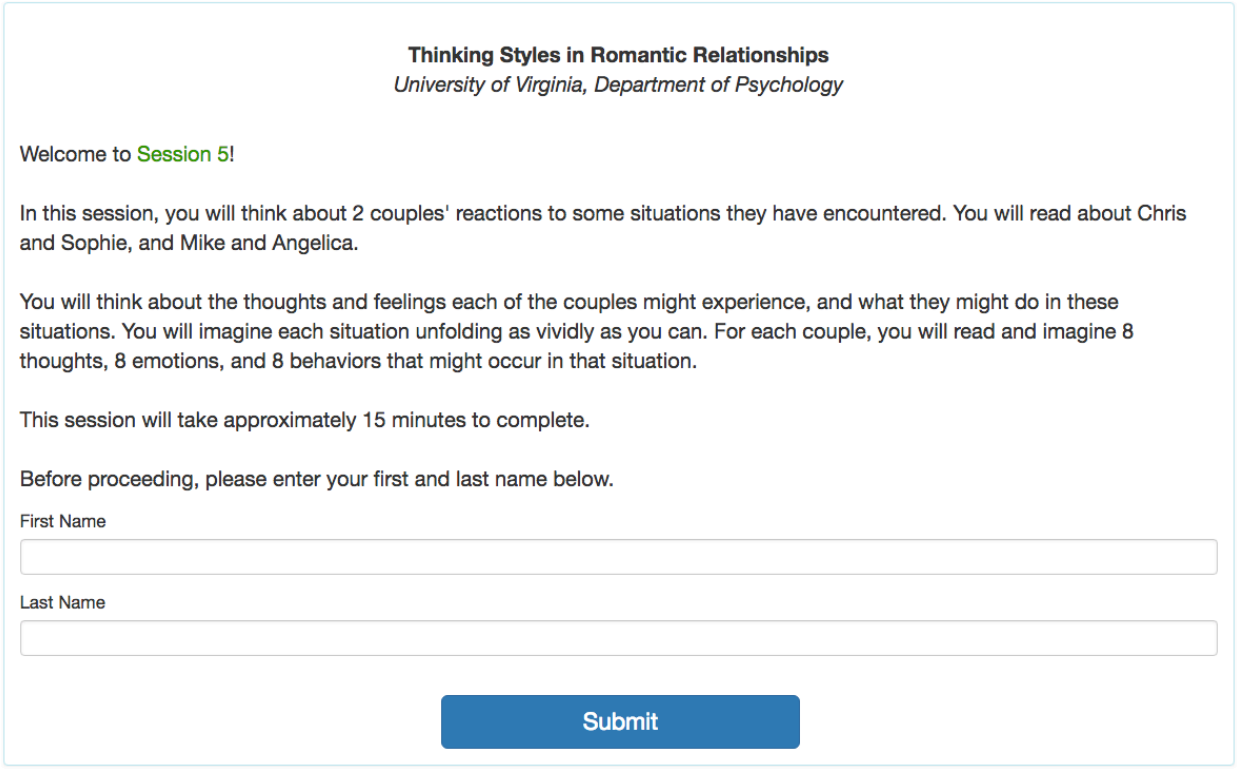
(d) Decliner. A reduction in scores between training vs. baseline.

Figure 8. Sample plots of participant relational flexibility over time.
Dashed lines: = 2-SD band; Solid line = Mean; Dotted lines = range.

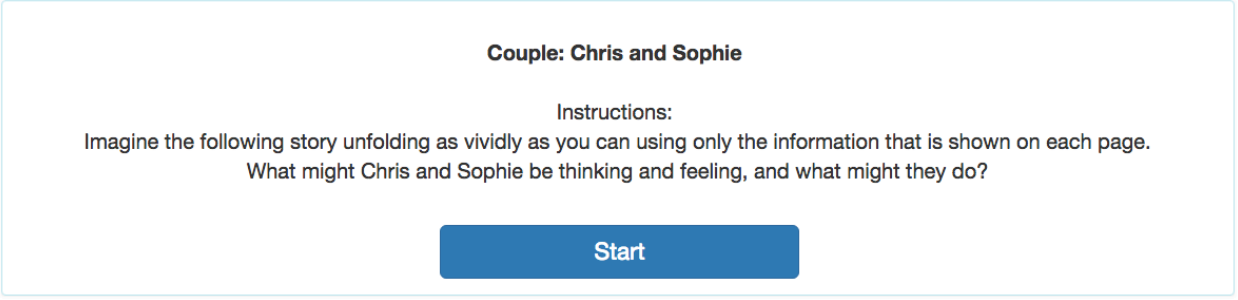
Appendix A. Sample Training Materials

Sample training scenario with Chris and Sophie for the "Communication" domain:

General instructions:



Instructions for specific scenario:



Presentation of scenario:

Chris makes dinner for Sophie and himself. They eat in silence.

Click **Start** when you are ready to start imagining.

Start

Instructions following initial imagery:

Next, you'll be given information about different ways Chris and Sophie might be thinking or feeling, and different things they might do in the situation.

Fill in the missing letter. Then, you will be asked to imagine as vividly as you can the story unfolding using only the information that is given on that page. Afterwards, you will answer a simple question about what you read.

Next

Presentation of sentence stem:

You know that:
Chris makes dinner for Sophie and himself. They eat in silence.

What might Chris be thinking?

Chris thinks "The meal I made tastes awf_l."

Type the missing letter below:

Next

Sentence stems for Attributions-Block: (Order of attributions randomized. Last item presented will always be a positive attribution)

From Chris' perspective

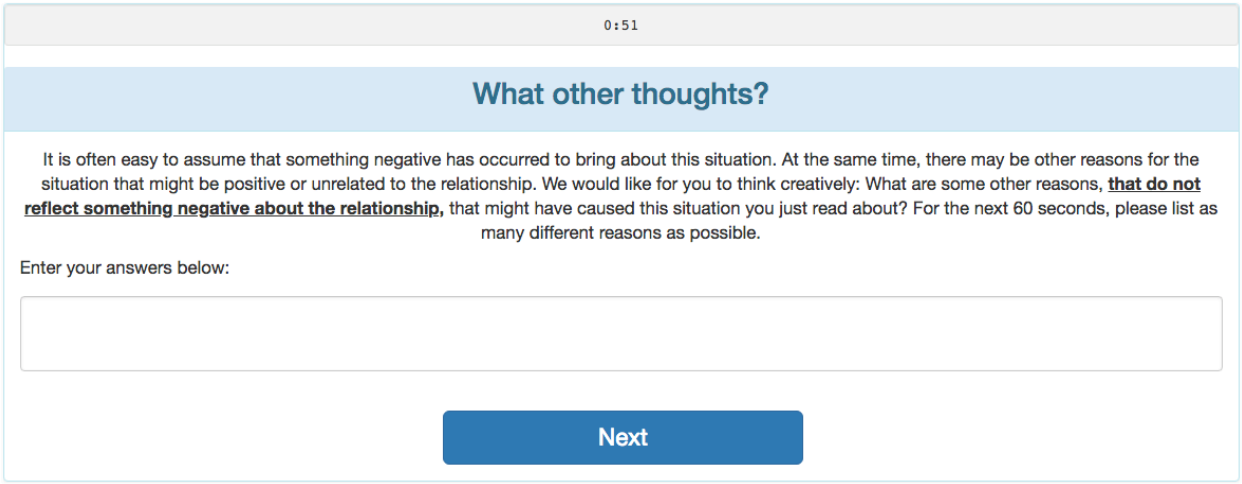
Attribution #1: Related to self (e.g., you think that the cause of the event is something directly related to you) (negative)	Chris thinks "The meal I made tastes awful." Did Chris like the taste of the food he made?
Attribution #2: Related to self (positive)	Chris thinks "I understand Sophie's food preferences very well." Did Chris think he understood what Sophie liked to eat?
Attribution #3: Related to partner (e.g., it was something about him/her, such as the kind of person he/she is) (negative)	Chris thinks "Sophie has never liked my cooking." Did Chris think that Sophie enjoys the meals he makes?
Attribution #4: Related to partner (positive)	Chris thinks "Sophie is quietly enjoying her meal." Did Chris think Sophie disliked the meal?
Attribution #5: Related to events outside the relationship (negative)	Chris thinks "Sophie is feeling ill and does not have much of an appetite." Did Chris think Sophie was feeling healthy?
Attribution #6: Related to events outside the relationship (positive)	Chris thinks "Sophie is thinking about the opportunity for promotion at work." Did Chris think Sophie was thinking about good news at work?
Attribution #7: Related to events in the relationship (negative)	Chris thinks "Sophie is still upset about our argument yesterday." Did Chris and Sophie have an argument yesterday?
Attribution #8: Related to events in the relationship (positive)	Chris thinks "This meal probably reminds Sophie of the wonderful dinner we had on our first date." Did Chris think they had a great meal together on their first date?

From Sophie's perspective:

Attribution #9: Related to self (negative)	Sophie thinks "I'm not contributing enough to the relationship." Did Sophie think she was contributing enough to the relationship?
Attribution #10: Related to self (positive)	Sophie thinks "Quietly savoring the meal is the best way to show my appreciation." Did Sophie think she was showing appreciation for the meal?
Attribution #11: Related to partner (e.g., it was something about him/her, such as the kind of person he/she is) (negative)	Sophie thinks "Chris is tired from making the meal." Did Sophie think Chris was energized from making dinner?
Attribution #12: Related to partner (positive)	Sophie thinks "Chris is really enjoying the taste of this food." Did Sophie think Chris disliked the meal?
Attribution #13: Related to events outside the	Sophie thinks "Chris is worried about his presentation tomorrow."

relationship (negative)	Did Sophie think Chris was confident about his presentation tomorrow?
Attribution #14: Related to events outside the relationship (positive)	Sophie thinks "The peaceful evening is perfect for a quiet dinner together."
	Did Sophie think the quiet evening was perfect?
Attribution #15: Related to events in the relationship (negative)	Sophie thinks "We never have in-depth conversations about anything anymore."
	Did Sophie think they frequently had in-depth conversations?
Attribution #16: Related to events in the relationship (positive)	Sophie thinks "We feel comfortable with silence and continue to enjoy each other's company."
	Did Sophie think that silence was uncomfortable?

Question to encourage flexible thinking after "Attributions-block". Participants will be provided with an empty text box where they list additional reasons that might be behind the ambiguous situation presented:

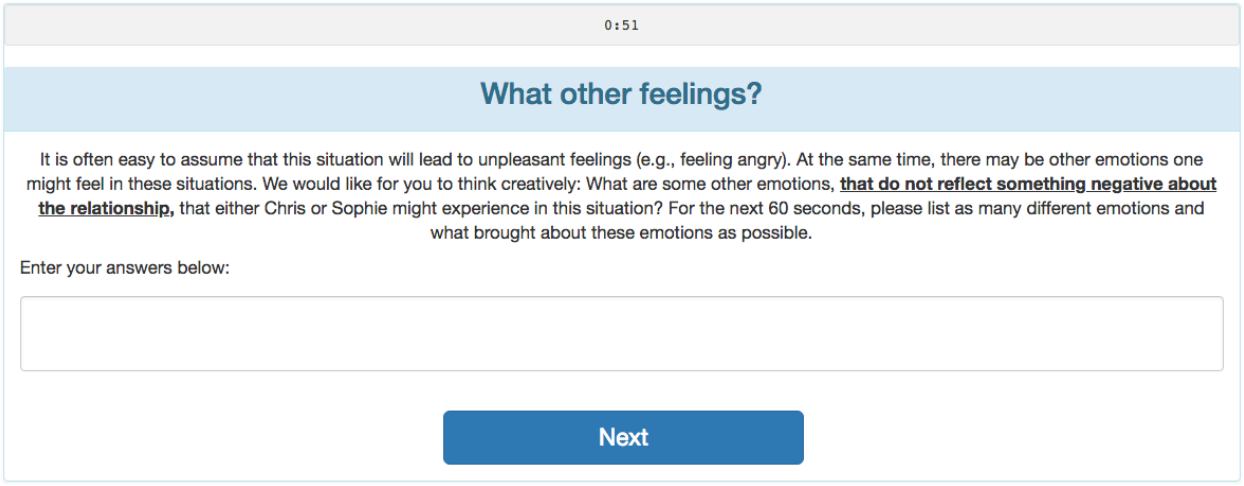


Emotions-Block: (Order of emotions randomized. Last item presented will always be a positive emotion)

From Chris's perspective:	
Emotion #1: An emotion you might feel in the situation (negative)	Chris is hurt by Sophie's silence.
	Did Chris feel encouraged by Sophie's silence?
Emotion #2: Another emotion you might feel in the situation (positive)	Chris is glad that Sophie appears to be enjoying the dinner.
	Did Chris think that Sophie was enjoying dinner?
Emotion #3: Another emotion you might feel in the situation (negative)	Chris is angry at Sophie for being so quiet.
	Did Chris like his partner's silence?
Emotion #4: Another emotion you might feel in the situation (positive)	Chris enjoys the meal in silence.
	Did Chris enjoy the meal in silence?
From Sophie's perspective:	
Emotion #5: An emotion your partner might	Sophie feels disappointed with the quality of the dinner.

feel (negative)	Was the quality of the meal satisfactory to Sophie?
Emotion #6: Another emotion your partner might feel (positive)	Sophie is pleasantly surprised that Chris made their dinner. Did Sophie like that Chris made their dinner?
Emotion #7: Another emotion your partner might feel (negative)	Sophie is annoyed at Chris for something he had done the previous day. Was Sophie pleased by Chris's actions the previous day?
Emotion #8: Another emotion your partner might feel (positive)	Sophie feels loved because Chris had volunteered to make dinner. Did Sophie like that Chris made dinner?

Question to encourage flexible thinking after "Emotions-block". Participants will be provided with an empty text box where they list additional emotions the couple might feel when experiencing the ambiguous situation presented.



Behavior Block: (Order of behaviors randomized. Last item presented will always be a positive behavior)

From Chris's perspective:	
Behavior #1: An action Chris might do in the situation (positive)	Chris suggests that they discuss what each of them like or dislike about the meal. Did Chris want to talk about what they thought about the meal?
Behavior #2: Another action Chris might do in the situation (negative)	Chris makes a sarcastic comment at Sophie for being unappreciative of his efforts to make dinner. Did Chris say something sarcastic to Sophie?
Behavior #3: Another action Chris might do in the situation (positive)	Chris lets Sophie know that her opinions are important to him. Did Chris value his partner's opinions?
Behavior #4: Another action Chris might do in the situation (positive)	Chris shows concern by asking Sophie if she is really tired. Did Chris express concern for Sophie?
From Sophie's perspective:	
Behavior #5: An action Sophie might do in	After dinner, Sophie expresses appreciation for Chris's efforts in making dinner.

the situation (positive)

Did Sophie remain silent after dinner?

Behavior #6: Sophie avoids making eye contact throughout dinner.

Another action Sophie might do in the situation (negative)

Did Sophie make eye contact with Chris during dinner?

Behavior #7: Sophie asks if Chris would like to chat about what's on his mind.

Another action Sophie might do in the situation (positive)

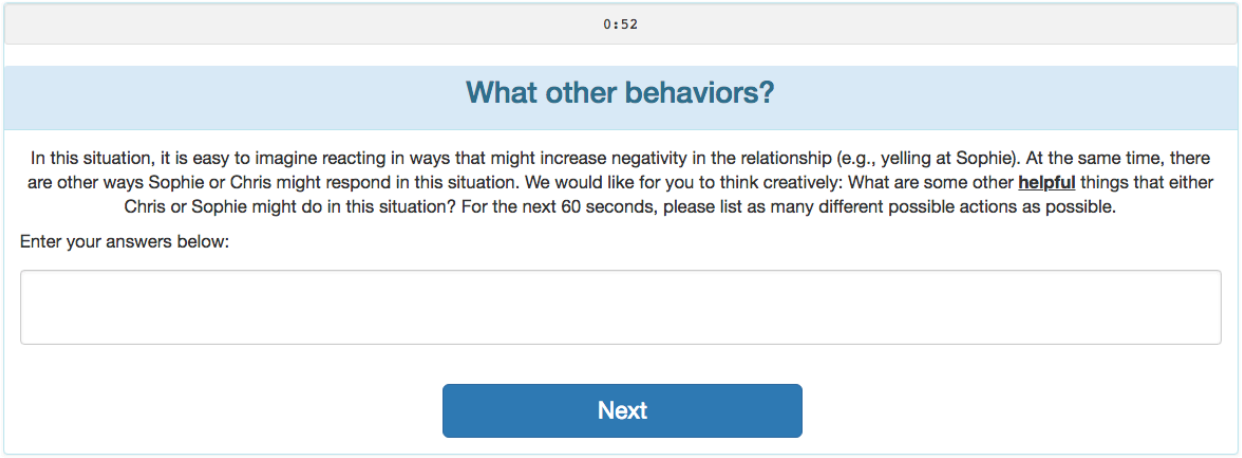
Would Sophie like to know what Chris was thinking?

Behavior #8: Sophie offers to help with making dinner tomorrow.

Another action Sophie might do in the situation (positive)

Would Sophie like to make dinner with Chris tomorrow?

Question to encourage flexible thinking after "Behaviors-block". Participants will be provided with an empty text box where they list additional positive behaviors the couple might do in the situation presented.



Appendix B. Example of CBM-FlexC Training Materials Assessment Protocol

Example Scenario evaluation form
(Participants will answer these questions for each scenario)

Scenario: [*Name of scenario*]

Problem Domain List: [*Communication, Sex/Intimacy, Children/other family (e.g., in-laws), Finances, Division of labor (e.g., household chores, childcare)*]

(Please circle the appropriate number)

1. **How relevant is the scenario to relationship-specific problem domain?**
   Not relevant at all 1 2 3 4 5 Extremely relevant

2. **How ambiguous (with respect to being able to interpret what is going on in the scenario) do you think the situation is?**
   Not ambiguous at all 1 2 3 4 5 Extremely ambiguous

3. **How anxious would you be in a similar situation?**
   Not anxious at all 1 2 3 4 5 Extremely anxious

Note: Participants were also asked to note whether the wording is clear and the content seems plausible.

Appendix C. Sample Training Schedule

Example: Participant X is in a heterosexual relationship. Before starting session 1, she selects communication, division of labor, and finances as the domains she needs help with. She is assigned, by the computer program, to follow the fictitious couples Chris and Sophie, and Mike and Angelica. Her training will be scheduled as follows:

Fictional couple:	**Chris and Sophie**	**Mike and Angelica**
Session 1:	Communication (domain #1)	Finances (domain #3)
Session 2:	Division of labor (domain #2)	Communication (domain #1)
Session 3:	Finances (domain #3)	Division of labor (domain #2)
Session 4:	Communication (domain #1)	Division of labor (domain #2)
Session 5:	Division of labor (domain #2)	Finances (domain #3)
Session 6:	Finances (domain #3)	Communication (domain #1)

Appendix D. Measurement Plan

Study 1a

- Online screening questionnaire
- Aggregated Relational Flexibility (ARF; containing 25 shortlisted items)
- Cognitive Flexibility Inventory (CFI-20; construct validity; as control for incremental validity)
- Interpersonal Reactivity Index for Couples (IRIC; construct validity; as control for incremental validity)
- Relationship Attribution Measure (RAM; predictive validity; as criterion variable for incremental validity)
- Couples Satisfaction Index (CSI-4; predictive validity; as criterion variable for incremental validity)
- Depression, Anxiety, and Stress Scales (DASS-21; discriminant validity)

Study 1b

Initial Test:

- Online screening questionnaire
- Aggregated Relational Flexibility (ARF with 14 items; 10 items on ARF-PT subscale and 4 items on ARF-NR subscale)
- Cognitive Flexibility Inventory (CFI-10; construct validity; as control for incremental validity)
- Interpersonal Reactivity Index for Couples (IRIC; construct validity; as control for incremental validity)
- Couples Satisfaction Index (CSI-4; predictive validity; as criterion variable for incremental validity)
- Patient Health Questionnaire for Depression and Anxiety (PHQ-4; discriminant validity)

Retest with subset of Ps:

- Aggregated Relational Flexibility (ARF with 14 items; 10 items on ARF-PT subscale and 4 items on ARF-NR subscale)

Study 2

- Qualitative interview
- Assessment protocol for evaluating CBM-FlexC training materials

Study 3

Screening

- Online screening questionnaire
- Aggregated Relational Flexibility (8-item ARF)
- Couples Satisfaction Index (CSI-4)
- Suicidal Behaviors Questionnaire-Revised (SBQ-R; administered during phone screening interview)
- Hurt, Insult, Threaten, and Scream Questionnaire (HITS; administered during phone screening interview)

Baseline Assessment #1, and Assessments after CBM-FlexC training sessions 2, 4, and 6, and One-Month Follow-Up Assessment

- Couples Satisfaction Index (CSI-4)
- Cognitive Flexibility Inventory (CFI-10)
- Aggregated Relational Flexibility (ARF)
- Differentiation of Self Index (DSI; only at Baseline #3, CBM-FlexC #6, and One-month followup)
- Patient Health Questionnaire for Depression and Anxiety (PHQ-4)
- Exit interview (only at one-month followup)

Baseline assessments #2 and #3

- Aggregated Relational Flexibility (ARF)

Appendix E. Questionnaires

Studies 1 and 3: Online screening questionnaire

Note: Study 1's online screening questionnaire will only include questions 1 to 3. Study 3's screening questionnaire will include all questions listed below, and also include suicide ideation and intimate partner violence assessments (using Suicidal Behaviors Questionnaire-Revised; Osman et al., 2001; and Hurt, Insult, Threat, Scream Questionnaire; Sherin et al., 1998).

1. Relationship status (pick all that apply):

A. Single
B. In a relationship
C. Married
D. Engaged
E. Civil Union
F. Domestic partnership
G. Open relationship
H. Separated
I. Divorced
J. Widowed

(items #2 and #3 only appear if answers to Q1 contains: B, C, D, E, F, G)
2. How long have you been with your current romantic partner?

A. Less than 3 months
B. 3-6 months
C. 6-12 months
D. 1-2 years
E. 3-5 years
F. 5-10 years
G. More than 10 years

3. Are you in a long distance relationship?
Y / N

4. Are either you or your partner making concrete plans to separate/divorce?
Y / N

5. Are either you or your partner currently involved in an ongoing affair?
Y / N

6. Are you and your partner currently in couples therapy?
Y / N

7. Do you have regular access to a private, high-speed internet connection?
Y / N

Cognitive Flexibility Inventory

Note: Greyed items are those that were retained for CFI-10.

Instructions: Please use the scale below to indicate the extent to which you agree or disagree with the following statements.

Strongly disagree	Disagree	Somewhat disagree	Neutral	Somewhat agree	Agree	Strongly agree
1	2	3	4	5	6	7

1.	I am good at ‘‘sizing up’’ situations.	1	2	3	4	5	6	7
2.	I have a hard time making decisions when faced with difficult situations.	1	2	3	4	5	6	7
3.	I consider multiple options before making a decision.	1	2	3	4	5	6	7
4.	When I encounter difficult situations, I feel like I am losing control.	1	2	3	4	5	6	7
5.	I like to look at difficult situations from many different angles.	1	2	3	4	5	6	7
6.	I seek additional information not immediately available before attributing causes to behavior.	1	2	3	4	5	6	7
7.	When encountering difficult situations, I become so stressed that I cannot think of a way to resolve the situation.	1	2	3	4	5	6	7
8.	I try to think about things from another person’s point of view.	1	2	3	4	5	6	7
9.	I find it troublesome that there are so many different ways to deal with difficult situations.	1	2	3	4	5	6	7
10.	I am good at putting myself in others’ shoes.	1	2	3	4	5	6	7
11.	When I encounter difficult situations, I just don’t know what to do.	1	2	3	4	5	6	7
12.	It is important to look at difficult situations from many angles.	1	2	3	4	5	6	7
13.	When in difficult situations, I consider multiple options before deciding how to behave.	1	2	3	4	5	6	7
14.	I often look at a situation from different viewpoints.	1	2	3	4	5	6	7

15.	I am capable of overcoming the difficulties in life that I face.	1	2	3	4	5	6	7
16.	I consider all the available facts and information when attributing causes to behavior.	1	2	3	4	5	6	7
17.	I feel I have no power to change things in difficult situations.	1	2	3	4	5	6	7
18.	When I encounter difficult situations, I stop and try to think of several ways to resolve it.	1	2	3	4	5	6	7
19.	I can think of more than one way to resolve a difficult situation I'm confronted with.	1	2	3	4	5	6	7
20.	I consider multiple options before responding to difficult situations.	1	2	3	4	5	6	7

Interpersonal Reactivity Index for Couples

Instructions: The following statements inquire about your thoughts and feelings in a variety of situations occurring in your relationship with your partner. For each item, indicate how well it describes you by circling the appropriate number.

0 = Does not describe me well

4 = Describes me very well.

1.	I often have tender, concerned feelings for my partner when he/she is less fortunate than me.	0	1	2	3	4
2.	Sometimes I don't feel very sorry for my partner when he/she is having problems.	0	1	2	3	4
3.	I try to look at my partner's side of a disagreement before I make a decision.	0	1	2	3	4
4.	When I see my partner being taken advantage of, I feel kind of protective towards him/her.	0	1	2	3	4
5.	I sometimes try to understand my partner better by imagining how things look from his/her perspective.	0	1	2	3	4
6.	My partner's misfortunes do not usually disturb me a great deal.	0	1	2	3	4
7.	If I'm sure I'm right about something, I don't waste much time listening to my partner's arguments.	0	1	2	3	4
8.	When I see my partner being treated unfairly, I sometimes don't feel very much pity for him/her.	0	1	2	3	4
9.	I am often quite touched by things I see happen in my relationship.	0	1	2	3	4
10.	In my relationship, I believe that there are two sides to every question and try to look at them both.	0	1	2	3	4
11.	In my relationship with my partner, I would describe myself as a pretty soft-hearted person.	0	1	2	3	4
12.	When I'm upset at my partner, I usually try to "put myself in his/her shoes" for a while.	0	1	2	3	4
13.	Before criticizing my partner, I try to imagine how I would feel if I were in his/her place.	0	1	2	3	4

Short Relationship Attribution Measure (RAM)

Instructions: This questionnaire describes several things that your partner might do. Imagine your partner performing each behavior and then read the statements that follow it. Please circle the number that indicates how much you agree or disagree with each statement, using the rating scale below:

1	2	3	4	5	6
Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree

YOUR PARTNER CRITICIZES SOMETHING YOU SAY						
My partner's behavior was due to something about him/her (e.g., the type of person he/she is, his/her mood)	1	2	3	4	5	6
My partner's behavior was due to something about me (e.g., the type of person I am, the mood I was in)	1	2	3	4	5	6
The reason my partner criticized me is not likely to change	1	2	3	4	5	6
The reason my partner criticized me is something that affects other areas of our marriage	1	2	3	4	5	6
My partner criticized me on purpose rather than unintentionally	1	2	3	4	5	6
My partner's behavior was motivated by selfish rather than unselfish concerns	1	2	3	4	5	6
My partner deserves to be blamed for criticizing me	1	2	3	4	5	6
YOUR PARTNER BEGINS TO SPEND LESS TIME WITH YOU						
My partner's behavior was due to something about him/her (e.g., the type of person he/she is, his/her mood)	1	2	3	4	5	6
My partner's behavior was due to something about me (e.g., the type of person I am, the mood I was in)	1	2	3	4	5	6
The reason my partner is beginning to spend less time with me is not likely to change	1	2	3	4	5	6
The reason my partner is beginning to spend less time with me is something that affects other areas of our marriage	1	2	3	4	5	6
My partner is beginning to spend less time with me on purpose rather than unintentionally	1	2	3	4	5	6
My partner's behavior was motivated by selfish rather than unselfish concerns	1	2	3	4	5	6
My partner deserves to be blamed for beginning to spend less time with me	1	2	3	4	5	6
YOUR PARTNER DOES NOT PAY ATTENTION TO WHAT YOU ARE SAYING						
My partner's behavior was due to something about him/her (e.g., the type of person he/she is, his/her mood)	1	2	3	4	5	6
My partner's behavior was due to something about me (e.g., the type of person I am, the mood I was in)	1	2	3	4	5	6
The reason my partner did not pay attention to me is not likely to change	1	2	3	4	5	6
The reason my partner did not pay attention to me is something that	1	2	3	4	5	6

affects other areas of our marriage						
My partner did not pay attention to me on purpose rather than unintentionally	1	2	3	4	5	6
My partner's behavior was motivated by selfish rather than unselfish concerns	1	2	3	4	5	6
My partner deserves to be blamed for not paying attention to me	1	2	3	4	5	6
YOUR PARTNER IS COOL AND DISTANT						
My partner's behavior was due to something about him/her (e.g., the type of person he/she is, his/her mood)	1	2	3	4	5	6
My partner's behavior was due to something about me (e.g., the type of person I am, the mood I was in)	1	2	3	4	5	6
The reason my partner was cool and distant is not likely to change	1	2	3	4	5	6
The reason my partner was cool and distant is something that affects other areas of our marriage	1	2	3	4	5	6
My partner was cool and distant on purpose rather than unintentionally	1	2	3	4	5	6
My partner's behavior was motivated by selfish rather than unselfish concerns	1	2	3	4	5	6
My partner deserves to be blamed for being cool and distant	1	2	3	4	5	6

Couples Satisfaction Index-4 (CSI-4)

1. Please indicate the degree of happiness, all things considered, of your relationship.

Extremely Unhappy	Fairly Unhappy	A Little Unhappy	Happy	Very Happy	Extremely Happy	Perfect
0	1	2	3	4	5	6

	Not at all true	A little true	Somewhat true	Mostly true	Almost completely true	Completely true
2. I have a warm and comfortable relationship with my partner.	0	1	2	3	4	5

	Not at all	A little	Somewhat	Mostly	Almost completely	Completely
3. How rewarding is your relationship with your partner?	0	1	2	3	4	5
4. In general, how satisfied are you with your relationship?	0	1	2	3	4	5

Depression, Anxiety, and Stress Scales (DASS-21)

Instructions: Please read each statement and circle a number 0, 1, 2 or 3 that indicates how much the statement applied to you over the past week. There are no right or wrong answers. Do not spend too much time on any statement.

0 Did not apply to me at all
1 Applied to me to some degree, or some of the time
2 Applied to me to a considerable degree, or a good part of time
3 Applied to me very much, or most of the time

1	I found it hard to wind down - S	0	1	2	3
2	I was aware of dryness of my mouth - A	0	1	2	3
3	I couldn't seem to experience any positive feeling at all - D	0	1	2	3
4	I experienced breathing difficulty (eg, excessively rapid breathing, breathlessness in the absence of physical exertion) - A	0	1	2	3
5	I found it difficult to work up the initiative to do things - D	0	1	2	3
6	I tended to over-react to situations - S	0	1	2	3
7	I experienced trembling (eg, in the hands) - A	0	1	2	3
8	I felt that I was using a lot of nervous energy - S	0	1	2	3
9	I was worried about situations in which I might panic and make a fool of myself - A	0	1	2	3
10	I felt that I had nothing to look forward to - D	0	1	2	3
11	I found myself getting agitated - S	0	1	2	3
12	I found it difficult to relax - S	0	1	2	3
13	I felt down-hearted and blue - D	0	1	2	3
14	I was intolerant of anything that kept me from getting on with what I was doing - S	0	1	2	3
15	I felt I was close to panic - A	0	1	2	3
16	I was unable to become enthusiastic about anything - D	0	1	2	3
17	I felt I wasn't worth much as a person - D	0	1	2	3
18	I felt that I was rather touchy - S	0	1	2	3
19	I was aware of the action of my heart in the absence of physical exertion (eg, sense of heart rate increase, heart missing a beat) - A	0	1	2	3
20	I felt scared without any good reason - A	0	1	2	3
21	I felt that life was meaningless - D	0	1	2	3

Patient Health Questionnaire (PHQ-4)

Over the last 2 weeks, how often have you been bothered by the following problems?

	Not at all	Several days	More than half the days	Nearly every day
1. Feeling nervous, anxious or on edge	0	1	2	3
2. Not being able to stop or control worrying	0	1	2	3
3. Little interest or pleasure in doing things	0	1	2	3
4. Feeling down, depressed, or hopeless	0	1	2	3

Suicidal Behaviors Questionnaire-Revised (SBR-Q)

Please circle the number beside the statement or phrase that best applies to you.

1. Have you ever thought about or attempted to kill yourself? (circle only one)

1= Never
2= It was just a brief passing thought
3a= I have had a plan at least once to kill myself but did not try to do it
3b= I have had a plan at least once to kill myself and really wanted to die
4a= I have attempted to kill myself, but did not want to die
4b= I have attempted to kill myself, and really hoped to die

2. How often have you thought about killing yourself in the past year? (circle only one)

1= Never
2= Rarely (1 time)
3= Sometimes (2 times)
4= Often (3-4 times)
5= Very Often (5 or more times)

3. Have you ever told someone that you were going to commit suicide, or that you might do it? (circle only one)

1= No
2a= Yes, at one time, but did not really want to die
2b= Yes, at one time, and really wanted to do it
3a= Yes, more than once, but did not want to do it
3b= Yes, more than once, and really wanted to do it

4. How likely is it that you will attempt suicide someday? (Circle only one)

0= Never
1= No chance at all
2= Rather Unlikely
3= Unlikely
4= Likely
5= Rather Likely
6= Very Likely

Hurt, Insult, Threat, and Scream (HITS) Questionnaire

How often does your partner:

	Never	Rarely	Sometimes	Fairly Often	Frequently
1. Physically hurt you?	1	2	3	4	5
2. Insult or talk down to you?	1	2	3	4	5
3. Threaten you with harm?	1	2	3	4	5
4. Scream or curse at you?	1	2	3	4	5

Differentiation of Self Index (DSI)

Please read each of the following statements carefully and decide how much the statement is generally true of you. If you believe that an item does not pertain to you (e.g., one or both of you parents are deceased), please answer the item according to your best guess about what your thoughts and feelings would be in that situation. Responses are rated on a scale of 1 (not at all true of me) to 6 (very true of me).

1. People have remarked that I'm overly emotional.
2. I have difficulty expressing my feelings to people I care for.
3. I often feel inhibited around my family.
4. I tend to remain pretty calm even under stress.
5. I'm likely to smooth over or settle conflicts between two people whom I care about.
6. When someone close to me disappoints me, I withdraw from him or her for a time.
7. No matter what happens in my life, I know that I'll never lose my sense of who I am.
8. I tend to distance myself when people get too close to me.
9. It has been said (or could be said) of me that I am still very attached to my parent(s).
10. I wish that I weren't so emotional.
11. I usually do not change my behavior simply to please another person.
12. My partner could not tolerate it if I were to express to him or her my true feelings about some things.
13. Whenever there is a problem in my relationship, I'm anxious to get it settled right away.
14. At times my feelings get the best of me and I have trouble thinking clearly.
15. When I am having an argument with someone, I can separate my thoughts about the issue from feelings about the person.
16. I'm often uncomfortable when people get too close to me.
17. It's important for me to keep in touch with my parents regularly.
18. At times, I feel as if I'm riding an emotional roller coaster.
19. There's no point in getting upset about things I cannot change.
20. I'm concerned about losing my independence in intimate relationships.
21. I'm overly sensitive to criticism.
22. When my partner is away for too long, I feel like I am missing a part of me.
23. I'm fairly self-accepting.
24. I often feel that my partner wants too much from me.
25. I try to live up to my parents' expectations.
26. If I have an argument with my partner, I tend to think about it all day.
27. I am able to say no to others even when I feel pressured by them.
28. When one of my relationships becomes very intense, I feel the urge to run away from it.
29. Arguments with my parent(s) or sibling(s) can still make me feel awful.
30. If someone is upset with me, I can't seem to let it go easily.
31. I'm less concerned that others approve of me than I am about doing what I think is right.
32. I would never consider turning to any of my family members for emotional support.
33. I find myself thinking a lot about my relationship with my partner.
34. I'm very sensitive to being hurt by others.
35. My self-esteem really depends on how others think of me.
36. When I'm with my spouse or partner, I often feel smothered.
37. I worry about people close to me getting sick, hurt, or upset.
38. I often wonder about the kind of impression I create.
39. When things go wrong, talking about them usually makes it worse.
40. I feel things more intensely than others do.
41. I usually do what I believe is right regardless of what others say.
42. Our relationship might be better if my partner would give me the space I need.
43. I tend to feel pretty stable under stress.

Aggregated Relationship Flexibility (ARF) Measure (Items shortlisted for Study 1A; See Table 1 for retained items that comprise final ARF measure)

1. When my partner and I disagreed on something, I looked at the issue from many different angles.
2. I considered multiple options before making a decision relating to my relationship/partner.
3. There was only one way to resolve a difficult relationship issue I was confronted with. (-)
4. When I had a conflict with my partner, I realized that I could react in many possible ways.
5. When I am feeling indecisive about an issue in my relationship, I felt that there was only one way right way to respond. (-)
6. There was often only one solution worth considering for handling a problem in my relationship. (-)
7. I was aware that there could be many reasons when my partner was not responsive to me.
8. I sought additional information before coming to a conclusion about why my partner behaved a certain way.
9. I made sure I had all the facts before reacting to relationship-related issues.
10. When discussing important relationship issues with my partner, I could communicate an idea to him/her in different ways.
11. I adjusted easily to changes in my relationship.
12. I preferred to keep things unchanged in my relationship. (-)
13. I was able to take advice from my partner.
14. I reacted negatively to criticism from my partner. (-)
15. I did not like hearing opinions from my partner that go against my way of thinking. (-)
16. I reacted strongly when my partner contradicted me. (-)
17. When my partner and I had different opinions, I found it difficult to compromise. (-)
18. I thought about an issue from my partner's point of view before I made a decision.
19. I tried to understand my partner better by imagining how things look from his/her perspective.
20. When I was sure I was right about something, I did not need to listen to my partner’s arguments. (-)
21. I believed that there are two sides to every issue in my relationship and tried to look at them both.
22. When I was upset at my partner, I tried to “put myself in his/her shoes” for a while.
23. Before criticizing my partner, I tried to imagine how I would feel if I were in his/her place.
24. I was always able to understand my partner's thoughts/feelings without checking with him/her.
25. I felt that a relationship has to be perfect in order to work out right. (-)

Appendix G. Qualitative Exit Interview

(adapted from Reuland, 2014)

NOTE: The interview does not have to proceed according to this sequence of questions; rather it can be like a conversation. However, please make sure you've covered each of these points.

I. Introduction

"Before we begin, I would like to explain what we will be doing today. Please feel free to ask any questions."

Explain purpose of interview.

- "The purpose of this interview is to gather information about your opinion of the thinking styles training program you and your partner recently completed. The interview will last approximately 30 minutes.

Explain that their input is very valuable/important to project.

- "I really appreciate your willingness to help us out with this important project. Your opinions are very valuable. We will take your input very seriously in helping to revise the training program and develop a better program for improving flexible thinking in relationships."

II. Events during or since data collection

Significant life events happening during data collection

- Did your partner or your family experience any positive, but unusual life events during our program? This might be like the birth of a child or getting a promotion at work?
- Did your partner or your family experience any negative or challenging life events during our program? Such as sickness, or family transition such as moving to a different home, or separation or divorce of a close friend or relative?
- Did any of these events interfere with or influence your use of the computer program?
- Did you introduce or continue any other therapies during the course of the intervention? Psychotherapy or psychopharm?
- Did you discuss the intervention with your partner or other family members?

III. Questions about effectiveness

- What did you think about the computer-based program? What were your impressions of it?
- What did you find beneficial?
- What was not helpful?
  (If not clear from response above) Overall, how helpful would you say the program was on a scale of 0-10 where 0 is not helpful at all, and 10 is extremely helpful?
- Do you feel like anything's changed with you or your partner since you started the computer program?
- Have your interactions with your partner changed? If so, how does it seem different? Are you and your partner having more productive discussions about important issues? Do you find it easier to engage your partner? Are the ways you and your partner deal with disagreements different? Do you find it easier to take the perspective of your partner

during discussions? (Note: Try to get at perspective taking, ability to regulate emotions during conflicts, engagement in productive discussions, ways in thinking about conflicts/relationship issues)
- What types of situations have you noticed you or your partner behaving/thinking/feeling differently in? (Are changes specific to particular kinds of interactions/contexts?)
- Do you see your partner thinking and behaving differently (toward you or other people)?
    - If yes, Why do you think these changes occurred?
    - If yes, Do you think these changes will last?
- Did you have any negative experiences due to the computer program?

IV. Questions about validity
- What did you think of the scenarios in the program?
- To what extent did the sentences describe situations you encounter or things that you worry about?
- Were any of the scenarios hard to understand?
- Are there other important situations we should include in the program that were not covered?

V. Questions about utility and feasibility

Format of CBM program
- What was it like for you using the computer program?
- Was any part of it confusing or hard to use?
- How much did the web program keep your interest and attention?
- How did you feel about the way the web program looked?
- Was having 6 sessions good, or too many, too few?
- How was the length of each session?
- How was the frequency of the sessions; that is, how was logging on twice a week?
- What would you think about doing the computer program at a clinic instead of at home?

VI. Questions about Credibility
- Did you think a computer program like this could help your ability to cope and react to situations that are stressful in your relationship when you first heard about it?
- [If answer is no] What would have made it more convincing to you?

VII. Any other suggestions for improvement
- Suggestions for improvement
- How can we improve the computer program?
- How can we improve other aspects of the study?
- Do you have any feedback about the interviews and assessments?
- Do you have any feedback about any members of the research team?

VIII. Conclusion

Summary of feedback and wrap-up
- Provide a brief overall summary of major opinions expressed. Then say “Please let me know if anything I’ve said doesn’t quite seem right. I really want to be sure I understood what you were telling me. ”

- Ask for any other feedback: Is there anything else that you would like to share with me that you have not had the opportunity to?"

Thank participant for their participation

- "Once again, we want to express our thanks for sharing your ideas. As we said, we will take your opinions into careful consideration when further developing our programs. Thank you so much for your time. We really appreciate your willingness to help us out."

IX. Debriefing

- Administer debriefing.