Journal of Family Psychology
2015, Vol. 29, No. 6, 807–817

0893-3200/15/$12.00 http://dx.doi.org/10.1037/fam0000134

Predicting Long-Term Risk for Relationship Dissolution Using Nonparametric Conditional Survival Trees

Sören Kliem
Criminological Research Institute of Lower Saxony, Hannover, Germany

Sarah Weusthoff and Kurt Hahlweg
Technische Universität Braunschweig

Katherine J. W. Baucom and Brian R. Baucom
University of Utah

Identifying risk factors for divorce or separation is an important step in the prevention of negative individual outcomes and societal costs associated with relationship dissolution. Programs that aim to prevent relationship distress and dissolution typically focus on changing processes that occur during couple conflict, although the predictive ability of conflict-specific variables has not been examined in the context of other factors related to relationship dissolution. The authors examine whether emotional responding and communication during couple conflict predict relationship dissolution after controlling for overall relationship quality and individual well-being. Using nonparametric conditional survival trees, the study at hand simultaneously examined the predictive abilities of physiological (systolic and diastolic blood pressure, heart rate, cortisol) and behavioral (fundamental frequency; f_0) indices of emotional responding, as well as observationally coded positive and negative communication behavior, on long-term relationship stability after controlling for relationship satisfaction and symptoms of depression. One hundred thirty-six spouses were assessed after participating in a randomized clinical trial of a relationship distress prevention program as well as 11 years thereafter; 32.5% of the couples' relationships had dissolved by follow up. For men, the only significant predictor of relationship dissolution was cortisol change score ($p = .012$). For women, only f_0 range was a significant predictor of relationship dissolution ($p = .034$). These findings highlight the importance of emotional responding during couple conflict for long-term relationship stability.

Keywords: divorce, couple conflict, emotional responding, f0 range, psychophysiology

One in two marriages in the United States and more than one in three marriages in Germany ends in separation and/or divorce (Destatis, 2012; U.S. National Center for Health Statistics, 2010). Despite rising divorce rates (40% to 55% in developed countries in general; Hahlweg, Grawe-Gerber, & Baucom, 2010), adults across different cultures still regard intimate relationships as the most important in their lives (Buss, 2004). Conflict-related communication behavior and emotional responding have a well-established association with relationship quality and ultimate relationship dissolution (see Kiecolt-Glaser & Newton, 2001, for review). Relationship dissolution is associated with a host of negative outcomes for men and women alike, including higher levels of substance use, more symptoms of depression and anxiety, more health problems, and a greater risk of all-cause mortality (Amato, 2010; Sbarra, Law, & Portley, 2011). Consequences of relationship dissolution are not limited to the individuals in the relationship. Individuals growing up without continuously married parents tend of have worse emotional, social, behavioral, physical health, and academic outcomes, as well as poorer psychological well-being in adulthood and increased likelihood of dissatisfaction and dissolution in their own relationships and marriages (Amato, 2010). Based on the well-documented negative outcomes associated with relationship dissolution, a central aim of intimate relationship research is to better understand the factors associated with relationship dissolution so that it can be prevented. Although a number of nonconflict-related predictors of relationship dissolution have been identified (Karney & Bradbury, 1995), the aim of the current study is to identify conflict-specific predictors of relationship dissolution after controlling for nonconflict variables. We elect to focus on conflict-specific predictors because of their central role in couple relationship education (CRE) programs. Below we review couple conflict variables that may predict relationship dissolution, as well

This article was published Online First July 20, 2015.

Sören Kliem, Criminological Research Institute of Lower Saxony, Hannover, Germany; Sarah Weusthoff and Kurt Hahlweg, Department of Psychology, Technical University of Braunschweig; Katherine J. W. Baucom and Brian R. Baucom, Department of Psychology, University of Utah.

Sören Kliem and Sarah Weusthoff share first authorship.

Preparation of this article was supported in part by grants from the Deutsche Forschungsgemeinschaft (DFG; Ha 1400/4-1 and Ha 1400/16-1) awarded to Kurt Hahlweg as well as by a start-up grant funded by the University of Utah awarded to Brian R. Baucom. The randomized clinical trial was funded by grants from the DFG awarded to Kurt Hahlweg (DFG HA 1400/1-2).

Correspondence concerning this article should be addressed to Sören Kliem, Criminological Research Institute of Lower Saxony, Lützerodestraβe 9, 30161 Hannover, Germany. E-mail: soeren.kliem@kfn.de

as nonconflict individual and relationship factors that serve as control variables in the current study.

Emotional Responding During Couple Conflict

Greater emotional distress during couple conflict is a robust predictor of a host of problematic long-term outcomes including subsequent relationship distress and dissolution (Gottman & Levenson, 1992). Emotions are thought to comprise experiential, physiological, and behavioral components (Mauss, Levenson, McCarter, Wilhelm, & Gross, 2005). Consistent with this perspective, emotional responses are reflected in experiential, physiological, and behavioral indices of emotion. Both subjective and objective approaches to the measurement of emotional responding have been used in relationship research. Although research using subjective approaches has linked greater experiential indices of emotional distress during couple conflict with poorer concurrent and subsequent relationship quality (e.g., postdiscussion observational ratings completed by each partner; Bloch, Haase, & Levenson, 2014), the vast majority of research has focused on objective measures of emotional responses. A great deal of work has examined *physiological* indices of emotional responding such as cardiovascular (e.g., heart rate, blood pressure), electrodermal (e.g., skin conductance), and endocrine system (e.g., cortisol, epinephrine) activity. There is a well-replicated finding that greater physiological reactivity is associated with poorer relationship functioning (e.g., Robles & Kiecolt-Glaser, 2003) and increased likelihood of subsequent divorce (Kiecolt-Glaser et al. 2003), with recent meta-analytic findings demonstrating small but significant associations between lower cardiovascular reactivity during conflict and greater relationship quality ($r = -.13$), but no significant links between cortisol reactivity during conflict and marital quality (Robles, Slatcher, Trombello, & McGinn, 2014).

Physiological indices of emotional responding such as cardiovascular (e.g., heart rate, blood pressure), electrodermal (e.g., skin conductance), and endocrine (e.g., cortisol, epinephrine) reactivity, have been extensively studied during relationship conflict. Physiological reactivity may be influenced primarily by the sympathetic (SNS; e.g., skin conductance) or the parasympathetic (PNS; e.g., heart rate variability) branches of the autonomic nervous system, or by a balance of the functioning of both branches (e.g., heart rate, blood pressure) (Diamond & Otter-Henderson, 2007). There is ample evidence that greater reactivity in numerous physiological systems is associated with poorer relationship functioning (e.g., Robles & Kiecolt-Glaser, 2003) and increased likelihood of subsequent divorce (Kiecolt-Glaser et al., 2003), with recent meta-analytic findings demonstrating small but significant associations (e.g., $r = -.13$ across multiple measures of cardiovascular reactivity; Robles, Slatcher, Trombello, & McGinn, 2014). At present, it is not clear whether physiological reactivity driven by the SNS, the PNS, or the combination of the two is more impactful for long-term relationship outcomes.

Behavioral indices of emotional responding (e.g., facial expression, vocal expression) also tend to be related to concurrent and longitudinal relationship quality (e.g., Gottman & Levenson, 1992). Observer ratings of partners' affect have demonstrated links with concurrent and longitudinal relationship satisfaction (e.g., Bloch, Haase, & Levenson, 2014). Another behavioral measure that is linked with relationship outcomes in recent work is vocally encoded emotional arousal (i.e., fundamental frequency or f_0). f_0 is based on the vocal properties of human speech. During the first phase of speech production, air is released from the lungs and passes over the opening and closing vocal folds in the larynx. Over time, varying rates of opening and closing occur with higher rates of opening and closing being associated with higher perceived voice pitch and higher f_0 values (measured in cps or hertz [Hz]; see Weusthoff, Baucom, & Hahlweg, 2013a for a recent review). Because of heightened muscle tension, arousal associated with physical or emotional activation affects the voice via the somatic nervous system (Larsen et al., 2008), with emotion-related arousal demonstrating stronger effects on f_0 than physical activity (Johannes et al., 2007). Higher levels of individual arousal lead to higher f_0 values, perceivable as higher voice pitch (Juslin & Scherer, 2005). Higher levels of f_0 during couple conflict are associated with higher concurrent levels of physiological arousal measures including blood pressure, heart rate, and cortisol (Weusthoff et al., 2013a). In longitudinal analyses, higher levels of f_0 during couple conflict are associated with a number of deleterious couple outcomes including worse response to behavioral couple therapy in a 2-year follow up assessment (Baucom, Atkins, Simpson, & Christensen, 2009) and recall of fewer communication skills 11 years after participation in a CRE program (Baucom, Weusthoff, Atkins, & Hahlweg, 2012). In the current study we examine objective measures of both physiological (i.e., blood pressure, heart rate, cortisol) and behavioral (i.e., fundamental frequency) indices of emotional responding.

It is important that couple communication during conflict be considered alongside measures of spouses' emotional responses. For example, Bloch et al. (2014) found that decreases in experiential and behavioral indices of wives' emotional distress during conflict (i.e., "downregulation of negative emotion") is associated with greater concurrent relationship satisfaction in both husbands and wives, and greater increases in relationship satisfaction for wives over 13 years. Importantly, the longitudinal effects were mediated by constructive communication; wives who were able to more quickly decrease instances of emotional distress were able to more effectively use constructive communication skills to solve problems. This work highlights the importance of considering emotional and communication variables simultaneously when examining prediction of long-term outcomes.

Communication During Couple Conflict

There is a substantial body of work focused on aspects of couple communication that predict long-term relationship outcomes (Heyman, 2001). The most robust findings are documented in studies using (objective) observational coding systems rather than (subjective) self-report measures, although the general pattern of findings is similar across methodological approaches. Satisfied couples typically display more adaptive/positive aspects of communication and fewer maladaptive/negative aspects of communication during conflict relative to their distressed counterparts. In addition to these group differences, studies have typically found links between relationship quality and communication. In a recent meta-analysis, Woodin (2011) found significant associations between greater relationship quality and (a) lower levels of observed hostility, distress, and withdrawal, and (b) higher levels of observed intimacy and problem solving during conflict. Finally, longitudinal studies have demonstrated that couple communication during conflict is predictive of subsequent divorce across both self-report (e.g., Birditt, Brown, Orbuch, & McIlvane, 2010) and observational measures (e.g., Markman, Rhoades, Stanley, & Peterson, 2013). We use observational measures of positive and negative communication in the current study.

Other Factors Related to Relationship Dissolution

Relationship and individual distress are important to consider in the prediction of relationship dissolution. Although effect sizes are modest, couples with lower relationship satisfaction are more likely to divorce (Karney & Bradbury, 1995). In a longitudinal study of the most widely researched CRE program, the Prevention and Relationship Enhancement Program (PREP; Markman, Stanley, & Blumberg, 2010), couples with lower levels of premarital relationship satisfaction were more likely to ultimately divorce or to continue to experience relational distress over the course of 13 years (e.g., Clements, Stanley, & Markman, 2004).

Individual psychological distress—and depression in particular—is also linked with relationship outcomes (Whisman, 2001). The stress generation model of depression (Hammen, 1991) emerged from findings that individuals with current or past depressive episodes were more likely to experience interpersonal stressors. This model, in which depression and stressful events are thought to reciprocally influence one another, has been supported through empirical studies of both relationship distress and divorce (e.g., Hammen & Brennan, 2002). Findings have been replicated in men and women across the life span with mood disorders, as well as in individuals with greater (subsyndromal) symptoms of depression (Hammen, 2006). Consistent with this body of research documenting increased likelihood of interpersonal stressors in individuals with depression, depression appears to have a larger impact on relationship dissolution than most other psychological disorders. In a recent multinational population survey, major depression had the largest population attributable risk for divorce out of a number of Axis I psychological disorders including substance use, anxiety, impulse-control, and mood disorders (Breslau et al., 2011). In the current study we examine whether conflict-specific variables (i.e., emotional responding and communication) are predictive of relationship dissolution after controlling for relationship and individual distress known to be related to the longitudinal course of relationships.

Current Study

In the current study we extend previous research in a number of ways. First, we use a multimethod approach to the study of relationship functioning through our use of self-report, physiological, behavioral, and observational methods of assessment. Second, we use multiple measures of emotional responding (i.e., systolic and diastolic blood pressure, heart rate, salivary cortisol, and fundamental frequency) in combination with observational measures of communication behavior. Third, we predict a relationship outcome (i.e., relationship stability) 11 years after initial assessment. We hypothesize that higher levels of systolic and diastolic blood pressure, faster heart rate, greater salivary cortisol output, and higher f_0 and poorer communication (i.e., more negative and less positive communication behavior) during conflict will predict relationship stability over an 11-year period after controlling for self-reported relationship quality and individual partner distress (i.e., self-reported depressive symptoms).

Method

Participants and Procedure

Couples in the current study originally participated in a randomized controlled trial of a German-based behavioral relationship distress prevention program, Ein Partnerschaftliches Lernprogramm (EPL, "A Learning Program for Couples"; Kaiser, Hahlweg, Fehm-Wolfsdorf, & Groth, 1998). EPL is a behaviorally oriented relationship distress prevention program based on the Premarital Relationship Enhancement Program (PREP; Markman, Renick, Floyd, Stanley, & Clements, 1993) and focuses on learning and applying communication and problem-solving skills. Before the EPL training, as well as after the training, couples participated in 2.5-hr assessment sessions. During these sessions, couples were interviewed and completed questionnaires. They also participated in a 15-min videotaped problem-solving interaction task, discussing one of their most important relationship problem areas. In this study we used the assessment results from the post assessment interview session as prognostic variables. This procedure was chosen to control for possible influences of the EPL-training. A potential training effect likely is tapped by the postassessment measures already. Hence the prognosis of divorce risk via self-report, physiological, behavioral and observational methods should not be distorted by training participation. Approximately 11 years after baseline, participants were contacted again to determine relationship stability. The study was approved by Ethical Review Board of the German Association of Psychology (DGPs).

The present analyses include 68 couples (N = 136 partners) from the original sample of 81 couples who completed the post-assessment session (n = 2: no information on timepoint of divorce/separation; n = 6: missing; n = 5: one partner of couple deceased). Details regarding the procedure used for assignment to treatment and for follow-up assessments are described in Hahlweg and Richter (2010). At baseline male partners were 39.9 years old (SD = 8.1, range: 25–64) and female partners were 37.1 years old (SD = 7.7, range = 23–62). The majority of couples (80%) were married and the mean duration of marriage or relationship was 10.9 years (SD = 8.1, range: 3–35). About 70% of the male and 60% of the female partners had at least high school education, about 90% of the male and 60% of the female partners were working and about 60% of the male and 70% of the female partners identified themselves as religious. All participants self-identified as Caucasian.

Measures

Relationship quality. The Partnership Questionnaire (PFB; Hahlweg, 1996) is a 30-item measure of relationship quality. It consists of three subscales (Quarreling, Tenderness, and Togetherness/Communication) which can be combined to yield a composite PFB score (Cronbach's alpha = .95). The PFB reliably discriminates distressed and nondistressed couples and sensitively monitors change resulting from couple therapy. The PFB correlates highly (r = .85; (Hahlweg, Klann & Hank, 1992) with the Dyadic Adjustment Scale (DAS; Spanier, 1976). A score below the cutoff of 54 is indicative of low relationship quality.

Communication behavior. Interactional behaviors were assessed based on couples' videotaped discussions of a relationship problem using the Kategoriensystem fuer partnerschaftliche Interaktion (KPI; Interactional Coding System; Hahlweg, 2005). Sum scores of Verbal Positive (Self-disclosure, Positive Solution, Acceptance, Agreement) and Verbal Negative (Criticism, Negative Solution, Justification, Disagreement) were calculated. All of the content categories also received a nonverbal rating as positive, negative or neutral. Cohen's Kappa coefficients were .78 ($p <$ .0001) for the verbal and .64 for the nonverbal codes ($p <$ .0001).

Depression. The German version of the Center for Epidemiological Studies-Depression Scale (CES-D; Radloff, 1977) was used to measure depression (Hautzinger & Bailer, 1992). It contains 20 items, was developed for epidemiological purposes, and has an internal consistency of .90. Raw scores greater than 16 are used as a cut-off for clinical depression.

Emotional Responding

Physiological indices. Several standardized protocols were used for measurement of cardiac and endocrine system variables. Systolic (SBP) and diastolic blood pressure (DBP), heart rate (HR), and cortisol output were each assessed four times during each assessment point: baseline (10 min after start of assessment session), immediately before (60 min) and immediately after a couple-conflict discussion (80 min) as well as at the end of debriefing (110 min). All cardiac variables were recorded with automatic cuffs (auscultatory blood pressure monitor: type HESTIA OZ 100) for both partners simultaneously while sitting quietly in armchairs. Salivary cortisol was sampled by Salivetten (Sarstedt, Germany) using a cotton swab participants held in their mouth. Saliva samples were thawed and centrifuged at 3,000 U/min before being assayed. Until assay, they were stored frozen at −20°C. All samples of a couple were analyzed in the same assay. The lower detection limit of assay was 0.82 nmol/L (intraassay variability was below 5%, interassay variability below 10%) (see Fehm-Wolfsdorf, Groth, Kaiser, & Hahlweg, 1999, for further information).

For HR, SBP, and DBP conflict measures were created for each index by averaging the samples collected at 60 and 80 min (i.e., samples immediately before and immediately after the conflict discussion). For cortisol, we created a conflict measure by averaging the samples collected at 80 and 110 min (i.e., immediately after the conflict discussion and at debriefing), as changes in cortisol output take approximately 15 to 20 min to occur (Kirschbaum & Hellhammer, 1994).

To account for individual baseline differences in HR, SBP, DBP, and cortisol, we subsequently subtracted the baseline values (10 min after start of assessment session) from the conflict values to model the amount of change during the discussion relative to resting baseline (change score = conflict values − baseline values). This procedure is a common approach to calculating physiological reactivity (e.g., Burt & Obradović, 2013).

Behavioral index. After continuously extracting f_0 from couples' problem-solving discussions sessions, all voice samples were bandpass-filtered to exclude any background noises before being analyzed using Praat (Boersma & Weenink, 2010). Afterward, several indices of f_0 (including mean, maximum, standard deviation, and range of f_0) can be statistically computed. Current methodological recommendations and empirical findings identify f_0 range as the cleanest measurement of vocally encoded emotional arousal of the computable indices (Juslin & Scherer, 2005); therefore, f_0 range (calculated by subtracting the minimum value of f_0 from the maximum value of f_0) was used in the current study. The typical range for male speakers is 75–150 Hz, for female speakers 150–300 Hz, though wider limits can be observed during highly aroused emotional states (Owren & Bachorowski, 2007). Higher levels of f_0 range indicate higher levels of emotional arousal. f_0 range was also used to enhance comparability with the other studies conducted this far in the couples area, with all findings being based on f_0 range (e.g., Baucom et al., 2009; Baucom et al., 2012).

Relationship stability at follow-up. Relationship stability (intact vs. divorced/separated) at 11-year follow-up was obtained by phone or mail. For couples that had divorced/separated the date of divorce/separation was collected.

Statistical Analysis

We tested associations between the output variable (relationship status at the 11-year follow-up) and the input variables from postassessment using the novel smooth Receiver Operator Characteristics (ROCs) Curves estimator based on log-concave density estimation (for a more detailed description, see Rufibach, 2012). The following input variables were used for both partners respectively: (a) relationship quality, (b) observationally coded verbal/nonverbal positive communication, (c) observationally coded verbal/nonverbal negative communication, (d) symptoms of depression, (e) f_0 range, (f) baseline SBP, (g) SBP change score, (h) baseline DBP, (i) DBP change score, (j) baseline HR, (k) HR change score, (l) baseline cortisol, and (m) cortisol change score. Results were provided as area under the curve (AUC), which can range from 0 to 1, and provide information similar to that yielded by effect size estimates (such as Cohen's d) making the predictive values of different measures comparable. According to Kraemer et al. (2003), standards for the AUC-values corresponding to those for Cohens d can be defined: a small effect size of d = 0.2 corresponds to AUC = .56; a medium effect size corresponds to AUC = .64, and a large effect size corresponds to AUC > .75. The advantage of using the AUC over more commonly used statistical measures of prediction (e.g., correlation coefficients or percent agreement) is that it is not constrained by base rates or selection ratios (Swets, 1986, 1988). Ninety-five percent confidence intervals were generated for the AUCs using 1,000 bootstrap resamples. We used the R packages "*pROC*" (Robin et al., 2011) and "*logcondens*" (Dümbgen & Rufibach, 2011) for this analysis.

In addition, the association between the input variables and relationship status at 11-year follow-up as well as the time period until divorce/separation was tested using nonparametric survival analysis (Surv-CTrees; also called recursive partitioning; Hothorn, Hornik, Strobl, & Zeileis, 2011; Hothorn, Hornik, & Zeileis, 2006), a form of Bayesian statistics. In Surv-CTrees, all parameters of interest can be analyzed simultaneously regardless of their distributions or interrelations with each other. Based on permutations, statistically relevant discriminatory power values for each variable of interest that are best able to discriminate between groups (in this case, divorced vs. not divorced/separated) are automatically identified. As most studies have tried to predict relationship stability using binary outcome variables (divorce—yes/no; e.g., Amato, 2010), the study at hand uses a different strategy. The dependent variable in this prediction study is composed of two parts: One is the time to event and the other is the event status (divorce), which records if the event of interest occurred or not. Surv-CTrees are a good alternative to (semi) parametric survival models (e.g., the Cox proportional hazard regression model; Bou-Hamad, Larocque, & Ben-Ameur, 2011), as parametric models have the drawback of forcing a specific link between covariates and responses and requiring interactions between covariates to be specified a priori. In contrast, survival trees are not subject to these limitations and offer greater flexibility (DeRose & Pallara, 1997). For example, certain types of interactions are detected automatically and do not have to be specified by the analyst.

Independence between any of the input variables and the response variable (not the independence between the input variables themselves) is assumed as a global hypothesis and is tested with the CTree algorithm. Group membership (EPL vs. non-EPL-takers) was included to control for any relevant influence. A tree is built by successively splitting the whole dataset into different groups, comparable to a stepwise regression, and nonsignificant predictors are excluded (see Kornhuber et al., 2013 for an example). When the algorithm stops, the hypothesis is rejected and the input variable with the strongest association with the response variable as well as the threshold c is included in the CTree. The strength of this association will be assessed by p value adjusted logrank statistics (Schumacher, Holländer, Schwarzer & Sauerbrei, 2001). Subsequently, the algorithm implements a binary split in the selected input variable and starts again from the beginning. For each subgroup, a weighted nonparametric Kaplan–Meier-Curve estimation of the survival function is carried out (Hothorn et al., 2006, 2011). Kaplan–Meier estimators offer the advantage of being computable for right-censored data (Kaplan & Meier, 1958).

All analyses were conducted with the freely available statistics software R (R Development Core Team, 2012). The "party" Package (Hothorn et al., 2011) was used to calculate the Surv-CTrees.

Outlier detection. Following the recommendation of Osborne and Overbay (2004), we searched for outliers in emotional arousal variables using interquartile range as recommended by Walfish (2006). A particularly useful variant of the interquartile range outlier method is called the median rule (Carling, 2000). The median rule is a modification of the classical boxplot outlier rule (e.g., Tukey, 1977) that performs better especially for small sample sizes (Carling, 2000). The frequency of outliers in physiological measures as well as in the f_0 values ranged between $n = 1$ (1.5%) and $n = 4$ (5.9%) for males and females, respectively. Outliers were handled as missing data in further analysis. We used the "*wls*" package (Wilcox, 2012) for this analysis.

Missing data. For some audio recordings ($n = 20$), audio quality was not sufficient for Praat analyses of vocal expressions, even after bandpass filtering. Because of the event's independence from the outcome variable relationship stability, we classified these cases as missing at random ("MAR"; McKnight, McKnight, Sidani, & Figueredo, 2007). No significant differences regarding input variables were detected comparing individuals with sufficient and with nonsufficient audio-data. The problem of missing data was corrected using the more novel nonparametric missing value imputation for mixed type data (missForest; Stekhoven & Bühlmann, 2012). Contrary to other potential ways of using imputation techniques (e.g., multiple imputation), missForest does not require any assumptions about distributional aspects of the data. The input variables for both partners, as well as age, gender, and relationship status at the 11-year follow-up were included in generating a complete data set. As shown in a recent simulation study (Stekhoven & Bühlmann, 2012), the missForest algorithm outperforms established imputation methods such as k-nearest neighbors imputation, or multivariate imputation using chained equations (van Buuren & Groothuis-Oudshoorn, 2011a, b). The R package "MissForest" (Stekhoven, 2012) was applied for this procedure.

Results

Twenty-two (32.5%) of the analyzed couples were divorced at the 11-year follow-up assessment. The mean amount of time between first-assessment and divorce was 4.7 years ($SD = 3.5$; range = 0.1–10 years). Table 1 presents the means and standard deviations of the input variables based on the imputed data set. At posttest, scores for women in verbal positive behaviors, f_0 range (for both analyses, $p = .018$), heart rate ($p = .044$), and f_0 range scores ($p < .001$) were significantly higher than for men, but men showed significantly higher levels of SBP ($p < .001$) and DBP ($p = .006$) than women. Furthermore, change scores of women in SBP were significantly higher than for men ($p = .008$). Low self-rated relationship quality, measured by PFBT scores below the cutoff of 54, was present for $n = 39$ (57.4%) of the male partners, and $n = 37$ (54.4%) of the female partners.

Predicting Long-Term Risk for Divorce

Table 2 displays the results of the ROC-analysis for each input variable and relationship status at the 11-year follow-up for female and male partners, respectively. For women, we found a significant positive association between postassessment f_0 values as well as for observationally coded verbal negative communication and risk for divorce at the 11year-follow up. For males, elevated levels of cortisol change scores as well as higher heart rate baseline scores at the postassessment were significantly related to a higher risk for divorce at the 11-year follow up. The AUCs of .66/.65 (association for females) and .65 (association for males) correspond to medium effect sizes (i.e., AUCs > .64). Furthermore, we found small but nonsignificant effect sizes (i.e., .64 ≥ AUCs > .56) for females' verbal positive (AUC = .59), verbal negative (AUC = .62), and nonverbal positive (AUC = .63) behaviors, $SBP_{change\text{-}score}$ (AUC = .62) as well as for self-rated relationship scores assessed via the PFB (AUC = .57). In addition, we found small but

Table 1
Means, Standard Deviations, and Gender Differences in Input Variables at Posttest

Input variables	Female partners (n = 68) *M*	Female partners (n = 68) *SD*	Male partners (n = 68) *M*	Male partners (n = 68) *SD*	*t*	*df*	*p*
KPI verbal positive	30.48	12.06	25.68	11.27	2.40	134	.018
KPI verbal negative	30.59	14.01	29.66	13.56	0.39	134	.695
KPI nonverbal positive	25.66	20.23	20.71	15.08	1.62	134	.108
KPI nonverbal negative	24.98	18.22	24.58	18.21	.013	134	.898
PFB	49.42	13.81	50.22	13.83	−0.34	134	.736
CES-D	13.68	7.98	11.39	6.49	1.84	134	.069
f_0 range scores	44.23	7.36	28.34	11.07	9.86	134	<.001
$SBP_{baseline}$	120.23	11.52	129.01	11.78	−4.40	134	<.001
$DBP_{baseline}$	74.49	9.10	79.12	10.19	−2.80	134	.006
$HR_{baseline}$	76.63	10.25	72.98	10.73	2.03	134	.044
$Cortisol_{baseline}$	0.12	0.08	0.14	0.10	−1.23	134	.222
$SBP_{change\text{-}score}$	1.38	6.91	−1.52	5.58	2.70	134	.008
$DBP_{change\text{-}score}$	2.41	4.93	1.18	4.61	1.50	134	.136
$HR_{change\text{-}score}$	−3.69	7.62	−4.15	4.25	0.44	134	.664
$Cortisol_{change\text{-}score}$	−0.28	0.05	−0.04	0.09	0.58	134	.560

Note. KPI = Kategoriensystem fuer partnerschaftliche Interaktion; PFB = partnership questionnaire; CES-D = Center for Epidemiological Studies Depression Scale; f_0 = fundamental frequency; SBP = systolic blood pressure; DBP = diastolic blood pressure; HR = heart rate.

nonsignificant effect sizes for males' $DBP_{baseline}$ (AUC = .64) and $SBP_{baseline}$ (AUC = .61) were detected.

We ran additional analyses using SurveC-Trees, with relationship stability and time period until divorce/separation as outcome variables. Results indicated a significant association between higher f_0 range and increased risk for divorce for female partners (p = .034; see Figure 1, left). In the male partners we found a significant association between cortisol change scores and increased risk for divorce (p = .012; see Figure 1, right). No other significant effects emerged in either analysis.

Table 2
AUCs Between Input Variables and Relationship Status at 11-Year Follow-Up for Male and Female Partners

Input variables	Female partners (n = 68) AUC	Female partners (n = 68) 95% CI[b]	Male partners (n = 68) AUC	Male partners (n = 68) 95% CI[b]
f0 range scores	**0.66**	[**0.55, 0.79**]	0.50	[0.36, 0.61]
KPI verbal positive[a]	0.59	[0.43, 0.74]	0.55	[0.39, 0.66]
KPI verbal negative	0.62	[0.48, 0.74]	0.55	[0.42, 0.67]
KPI nonverbal positive[a]	0.63	[0.50, 0.78]	0.44	[0.26, 0.61]
KPI nonverbal negative	**0.65**	[**0.52, 0.77**]	0.53	[0.38, 0.67]
PFB[a]	0.57	[0.39, 0.72]	0.47	[0.35, 0.66]
CES-D	0.51	[0.36, 0.65]	0.51	[0.38, 0.65]
$SBP_{baseline}$	0.37	[0.25, 0.54]	0.61	[0.45, 0.77]
$DBP_{baseline}$	0.46	[0.32, 0.59]	0.64	[0.48, 0.78]
$HR_{baseline}$	0.56	[0.39, 0.71]	**0.65**	[**0.51, 0.75**]
$Cortisol_{baseline}$	0.52	[0.37, 0.63]	0.46	[0.32, 0.60]
$SBP_{change\text{-}score}$	0.62	[0.46, 0.75]	0.52	[0.38, 0.65]
$DBP_{change\text{-}score}$	0.45	[0.29, 0.59]	0.50	[0.39, 0.62]
$HR_{change\text{-}score}$	0.47	[0.32, 0.62]	0.46	[0.33, 0.58]
$Cortisol_{change\text{-}score}$	0.49	[0.32, 0.61]	**0.65**	[**0.54, 0.78**]

Note. AUC = area under curve; KPI = Kategoriensystem fuer partnerschaftliche Interaktion; PFB = partnership questionnaire; CES-D = Center for Epidemiological Studies; Depression Scale; f_0 = fundamental frequency; SBP = systolic blood pressure; DBP = diastolic blood pressure; HR = heart rate. Values shown as bold are significant.
[a] The response variable was recoded for these analyses (separation/divorce = 0; still together at 11-year follow-up = 1), because higher scores on this measure indicate a lower risk for separation/divorce. [b] Bootstrapping 95% confidence intervals based on 1,000 bootstrap resamples. Confidence intervals not containing AUC = .5 are interpreted as significant (in bold).

Discussion

The primary aim of the current study was to examine depressive symptoms, psychophysiology, observationally coded behavior, vocally encoded emotional arousal (f_0 range), and initial relationship quality as predictors of relationship stability over 11 years. Consistent with hypotheses, greater wife f_0 range and higher husband cortisol output during problem-solving discussions were associated with increased longitudinal risk for relationship dissolution, and these associations did not vary as a function of whether couples received a couple-relationship enhancement intervention or not. After including these variables in the CTrees, no interaction effect with any of the other input variables could be detected. These findings are consistent with earlier work on predictors of long-term relationship stability in identifying conflict-related emotional responding as a particularly robust marker of a relationship's long-term viability.

Emotional Responding as a Predictor of Long-Term Relationship Stability

The associations between greater wife f_0 range, higher husband cortisol output, and increased longitudinal risk for relationship dissolution add additional empirical support to existing evidence of the predictive power of conflict-related emotional responding. Gottman and Levenson (1992, 1999, 2002) conducted a series of studies examining the utility of numerous self-report measures of relationship and individual functioning, observational coding of conflict behaviors, and physiological responding during couple conflict for predicting relationship outcomes over a 14-year period in a sample of community couples. Across these studies, higher

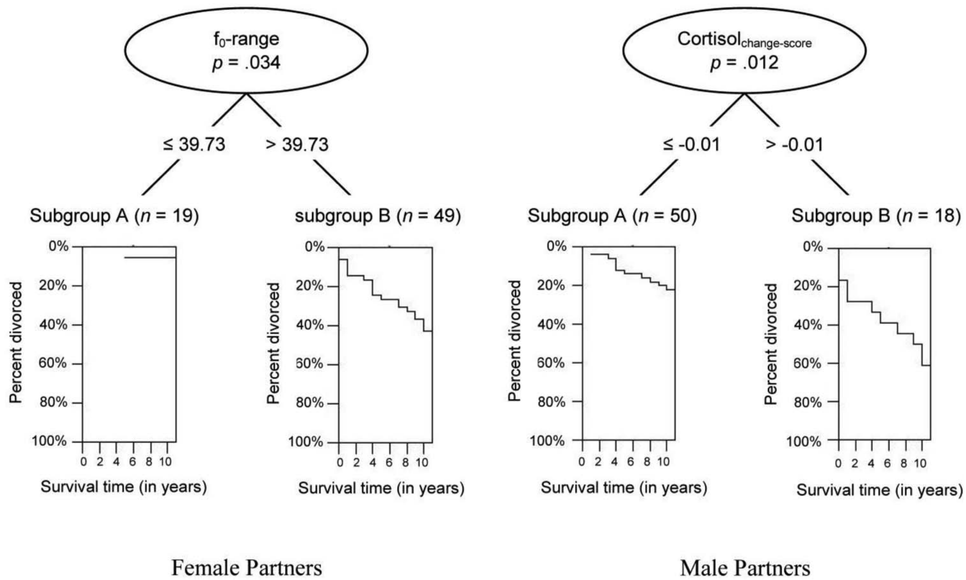

Figure 1. The risk for divorce in a nonparametric conditional inference trees.

levels of observationally coded negative communication behaviors, lower levels of observationally coded positive behaviors, and higher levels of physiological reactivity[1] were predictive of longitudinal declines in relationship satisfaction and increased risk for relationship dissolution. However, higher levels of physiological reactivity were the strongest predictor of 14-year risk for relationship dissolution; observationally coded positive and negative behaviors did not explain additional variance beyond that explained by physiological arousal. Convergent results are seen in the work of Kiecolt-Glaser et al. (2003) who found divorce over a 10-year period to be more robustly associated with neuroendocrine functioning during relationship conflict as well as during daily life than to observationally coded behavior during conflict or self-reported enduring emotional vulnerabilities.

Baucom et al. (2009) also examined a wide range of self-report, communication, and emotional variables as predictors of couples' responses to behaviorally based couple therapies two years after treatment termination. Bayesian model selection procedures (Austin & Tu, 2004; Raftery, 1995) were used to select an optimal subset of predictors; wives' conflict-related f_0 range emerged as one of four predictors selected from the 39 variables examined. Consistent with the results of the current study, greater wife f_0 range was associated with an increased likelihood of a poor response to couple therapy.[2]

Though there are numerous methodological and measurement differences between the current and previous studies, the substantive nature of the findings are highly consistent. Higher levels of conflict-related emotional distress and physiological reactivity are associated with increased likelihood of negative relationship outcomes, particularly over extended periods of time. Seemingly, physiological indicators are of particular importance for the long-term course and outcome of intimate relationships. The current study is consistent with Baucom et al. (2009) in finding this outcome for couples who have participated in a relationship intervention using f_0 range as a measure of vocally expressed emotional arousal as well as with Gottman and Levenson (1992, 1999, 2002) and Kiecolt-Glaser and colleagues (Kiecolt-Glaser et al., 2003) in finding this outcome for measures of physiological reactivity.

One methodological difference between the current study and Kiecolt-Glaser et al. (2003) that is particularly worthy of attention is that data in the current study were collected in the evening whereas data in Kiecolt-Glaser et al. (2003) were collected in the morning. Though there are a number of variables known to be associated with variability in diurnal rhythms (e.g., relationship functioning; Adam & Gunnar, 2001), cortisol levels typically increase in the morning and decrease in the evening (e.g., Weitzman et al., 1971). This typical diurnal pattern may have made it more difficult for Kiecolt-Glaser et al. (2003) to observe increases in cortisol levels caused by conflict discussions because of the concomitant increase due the typical diurnal rhythm and easier for the current study to observe increases in cortisol levels caused by conflict discussions because of the concomitant decrease due the typical diurnal rhythm.

The consistency of these findings across treatment-seeking and community couples in Germany and in the United States and across vocal and physiological measures of emotional responding suggests that spouses' emotional arousal during conflict is a vital element of couple functioning. This observation is by no means novel and has been made by numerous previous scholars (Karney & Bradbury, 1995; Markman et al., 2010). However, the results of the current study help to shed new light on this well-accepted conclusion and further deepen our understanding of the long-term

[1] Physiological responding was assessed using several measures, including heart rate, skin conductance, finger pulse amplitude, and pulse transmission time (Gottman & Levenson, 1999). Faster heart rate and greater skin conductance were most consistently associated with dysfunctional relationship outcomes, though there was some variability across studies.

[2] This association was strongest for couples that were moderately distressed prior to beginning treatment.

effects of psychophysiological indicators of emotional arousal in intimate relationships. Of particular note, f_0 range was a significant predictor of relationship dissolution for women and cortisol was a significant predictor for men. These gender differences in associations between behavioral and physiological measures of emotional responding were not hypothesized, and it is not immediate clear why such effects emerged. One possible interpretation of these findings is that they are related to which spouse was more actively pursuing change during the discussions. Evolutionary models of speech production propose that one function of emotional vocalization is to motivate behavioral responses from members of a social group (Juslin & Laukka, 2003). Consistent with this notion, Baucom and colleagues (2011) found that f_0 is associated with goal directed behaviors, such as demanding, but not with avoidant behaviors, such as withdrawing, during marital conflict. Thus it is likely that higher f_0 during marital conflict is an indication that a partner recognizes a problem in the relationship, is distressed about the problem, and wants the couple to take action to address the problem. As described in Weusthoff, Baucom, and Hahlweg (2013a), the discussions in the current study are characterized by significantly higher levels of female demand/male withdraw than male demand/female withdraw (i.e., demanding female partners and withdrawing male partners). This difference in demand/withdraw behavioral roles suggests that couples are likely discussing an issue where female partners are asking for change and male partners are resisting change. Bringing these factors together suggests that the highest risk for divorce likely occurs when female partners express a great deal of emotional arousal while pursuing change and male partner experience a strong physiological stress response while resisting change. However, demand/withdraw behavior was not included in the current manuscript so this possibility must be considered tentative until directly examined. Nevertheless, several studies (e.g., Margolin, Talovic, & Weinstein, 1983; Heyman, Hunt-Martorano, Malik, & Slep, 2009) demonstrated that women typically desire more relationship change than do men.

It is also important to consider these findings within the context of gender differences in earlier work on conflict-related emotional and physiological predictors of divorce. Gender differences emerged in Baucom et al. (2009), did not emerge in Kiecolt-Glaser et al. (2003), and were not tested in Gottman and Levenson (2002). Considering these findings as a whole, the most consistent and stable gender difference emerges for f_0. In both Baucom et al. (2009) and the current manuscript, only wives' f_0 was associated with divorce risk, and higher levels of wives' f_0 was associated with increased risk for divorce in both studies. In contrast, Kiecolt-Glaser et al. (2003) did not find any significant gender differences in associations between four stress hormones (ACTH, cortisol, epinephrine, and norepinephrine) and 10-year risk for divorce. Gottman and Levenson (2002) averaged physiological measures across spouses and analyzed couple level values. It is not clear why gender differences in physiological predictors of long-term divorce risk are inconsistent across existing studies and is an important area for future research.

Clinical Implications

Many couple therapies (Hahlweg et al., 2010) include interventions targeting aspects of couple's emotional lives. However, many of these interventions focus almost exclusively on controllable aspects of emotion, such as deliberate emotional expression (Christensen, 2010). The results of the current study suggest that involuntary aspects of emotional responding, such as cortisol output and f_0, may also be valuable targets for intervention. For example, exposure-based interventions have yet to be incorporated widely into couple therapies but could be helpful in reducing conditioned emotional responses during couple conflict (Christensen, 2010). Preliminary evidence suggests that couple therapy can interrupt conditioned emotional responding (Baucom, Atkins, & Christensen, 2010 as cited in Baucom & Atkins, 2013), but future research is needed to investigate this possibility.

Summary, Limitations, and Future Directions

The findings of the current study identify conflict-related emotional arousal as a robust predictor of long-term relationship outcomes over the longest follow-up period of any study of treatment-seeking couples to date. Combined with existing evidence linking conflict-related emotional arousal and physiological reactivity to long-term couple outcomes, these results suggest that involuntary emotional processes may be particularly important for long-term relationship stability.

Several limitations should be borne in mind when considering the results of the current study. First, although the sample size is comparable to other studies examining divorce prediction based on indices of emotional responding (e.g., Gottman & Levenson, 2002; Kiecolt-Glaser et al., 2003), f_0 range data was not evaluable for several participants. Well-accepted methods were used for imputing missing values, but it is still possible that this procedure impacted results. Second, the association between increased cortisol output during couple conflict and increased risk for divorce is inconsistent with some existing findings. Kiecolt-Glaser et al. (2003) did not find a significant association between cortisol and risk for divorce over 10 years, and Robles et al.'s (2014) meta-analysis concluded that there is no significant association between cortisol output and marital quality across four studies. The many methodological differences between studies of physiological and observationally coded predictors of long-term relationship health (e.g., no study includes the same set of variables as another) make it difficult to reconcile these inconsistent findings. Determining whether there is one, or a set of, physiological variable(s) with consistent predictive utility for long term relationship outcomes would be a valuable direction for future research. Third, the sample was composed entirely of Caucasian couples. Although the racial and ethnic composition of the sample is largely representative of the area from which couples were recruited, it is unclear how these findings would generalize to couples from other racial and ethnic backgrounds. Despite these limitations, the results of the current study identify two predictors of relationship outcomes for treatment-seeking couples over more than a decade of follow-up and suggest new avenues for refining and enhancing the effectiveness of couple therapies.

References

Adam, E. K., & Gunnar, M. R. (2001). Relationship functioning and home and work demands predict individual differences in diurnal cortisol patterns in women. *Psychoneuroendocrinology, 26,* 189–208. http://dx.doi.org/10.1016/S0306-4530(00)00045-7

Amato, P. R. (2010). Research on divorce: Continuing trends and new developments. *Journal of Marriage and Family, 72,* 650–666. http://dx.doi.org/10.1111/j.1741-3737.2010.00723.x

Austin, P. C., & Tu, J. V. (2004). Bootstrap methods for developing predictive models. *The American Statistician, 58,* 131–137. http://dx.doi.org/10.1198/0003130043277

Baucom, B. R., & Atkins, D. C. (2013). Polarization in marriage. In M. Fine & F. Fincham (Eds.), *Family theories: A content-based approach* (pp. 145–166). New York, NY: Routledge.

Baucom, B., Atkins, D., & Christensen, A. (2010). *Changes in vocally encoded emotional arousal in two behavioral couple therapies*. Poster presented at the triannual meeting of the World Congress of Behavioral and Cognitive Therapies, Boston, MA.

Baucom, B. R., Atkins, D. C., Eldridge, K., McFarland, P., Sevier, M., & Christensen, A. (2011). The language of demand/withdraw: Verbal and vocal expression in dyadic interactions. *Journal of Family Psychology, 25,* 570–580. http://dx.doi.org/10.1037/a0024064

Baucom, B., Atkins, D., Simpson, L., & Christensen, A. (2009). Prediction of response to treatment in a randomized clinical trial of couple therapy: A 2-year follow-up. *Journal of Consulting and Clinical Psychology, 77,* 160–173. http://dx.doi.org/10.1037/a0014405

Baucom, B. R., Weusthoff, S., Atkins, D. C., & Hahlweg, K. (2012). Greater emotional arousal predicts poorer long-term memory of communication skills in couples. *Behaviour Research and Therapy, 50,* 442–447. http://dx.doi.org/10.1016/j.brat.2012.03.010

Birditt, K. S., Brown, E., Orbuch, T. L., & McIlvane, J. M. (2010). Marital conflict behaviors and implications for divorce over 16 years. *Journal of Marriage and the Family, 72,* 1188–1204. http://dx.doi.org/10.1111/j.1741-3737.2010.00758.x

Bloch, L., Haase, C. M., & Levenson, R. W. (2014). Emotion regulation predicts marital satisfaction: More than a wives' tale. *Emotion, 14,* 130–144. http://dx.doi.org/10.1037/a0034272

Boersma, P., & Weenink, D. (2010). *Praat: Doing phonetics by computer* (Version 5.1.41) [Computer program]. Retrieved from http://www.praat.org/

Bou-Hamad, I., Larocque, D., & Ben-Ameur, H. (2011). A review of survival trees. *Statistics Surveys, 5,* 41–77. http://dx.doi.org/10.1214/09-SS047

Breslau, J., Miller, E., Jin, R., Sampson, N. A., Alonso, J., Andrade, L. H., . . . Kessler, R. C. (2011). A multinational study of mental disorders, marriage, and divorce. *Acta Psychiatrica Scandinavica, 124,* 474–486. http://dx.doi.org/10.1111/j.1600-0447.2011.01712.x

Burt, K. B., & Obradović, J. (2013). The construct of psychophysiological reactivity: Statistical and psychometric issues. *Developmental Review, 33,* 29–57. http://dx.doi.org/10.1016/j.dr.2012.10.002

Buss, D. M. (2004). *Evolutionary psychology: The new science of the mind*. Boston, MA: Pearson.

Carling, K. (2000). Resistant outlier rules and the non-Gaussian case. *Computational Statistics & Data Analysis, 33,* 249–258. http://dx.doi.org/10.1016/S0167-9473(99)00057-2

Christensen, A. (2010). A unified protocol for couple therapy. In K. Hahlweg, M. Grawe-Gerber, & D. Baucom (Eds.), *Enhancing couples: The shape of couple therapy to come* (pp. 33–46). Gottingen, Germany: Hogrefe Publishing.

Clements, M. L., Stanley, S. M., & Markman, H. J. (2004). Before they said "I do": Discriminating among marital outcomes over 13 years. *Journal of Marriage and Family, 66,* 613–626. http://dx.doi.org/10.1111/j.0022-2445.2004.00041.x

De Rose, A., & Pallara, A. (1997). Survival trees: An alternative non-parametric multivariate technique for life history analysis. *European Journal of Population, 13,* 223–241. http://dx.doi.org/10.1023/A:1005844818027

Destatis. (2012). *Scheidungen. Maβzahlen zur Ehescheidungen 2000 bis 2011*. [Divorce in Germany 2001–2011]. Retrieved from https://www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/Bevoelkerung/Ehescheidungen/Tabellen/MasszahlenEhescheidungen.html

Diamond, L. M., & Otter-Henderson, K. D. (2007). Physiological measures. In R. W. Robbins (Ed.), *Handbook of research methods in personality psychology* (pp. 370–388). New York, NY: Guilford Press Publications.

Dümbgen, L., & Rufibach, K. (2011). logcondens: Computations related to univariate log-concave density estimation. *Journal of Statistical Software, 39,* 1–28.

Fehm-Wolfsdorf, G., Groth, T., Kaiser, A., & Hahlweg, K. (1999). Cortisol responses to marital conflict depend on marital interaction quality. *International Journal of Behavioral Medicine, 6,* 207–227. http://dx.doi.org/10.1207/s15327558ijbm0603_1

Gottman, J. M., & Levenson, R. W. (1992). Marital processes predictive of later dissolution: Behavior, physiology, and health. *Journal of Personality and Social Psychology, 63,* 221–233. http://dx.doi.org/10.1037/0022-3514.63.2.221

Gottman, J. M., & Levenson, R. W. (1999). What predicts change in marital interaction over time? A study of alternative models. *Family Process, 38,* 143–158. http://dx.doi.org/10.1111/j.1545-5300.1999.00143.x

Gottman, J. M., & Levenson, R. W. (2002). A two-factor model for predicting when a couple will divorce: Exploratory analyses using 14-year longitudinal data. *Family Process, 41,* 83–96. http://dx.doi.org/10.1111/j.1545-5300.2002.40102000083.x

Hahlweg, K. (1996). *Fragebogen zur Partnerschaftsdiagnostik (FPD)* [Questionnaires for Partnership Diagnostics]. Goettingen, Germany: Hogrefe.

Hahlweg, K. (2005). Kategoriensystem für partnerschaftliche Interaktion (KPI): Interactional Coding System (ICS). In P. K. Kerig & D. H. Baucom (Eds.), *Couple observational coding systems* (pp. 127–142). Mahwah, NJ: Erlbaum.

Hahlweg, K., Grawe-Gerber, M., & Baucom, D. (2010). (Eds.). *Enhancing couples: The shape of couple therapy to come*. Gottingen, Germany: Hogrefe Publishing.

Hahlweg, K., Klann, N., & Hank, G. (1992). Zur Erfassung der Ehequalität: Ein Vergleich der "Dyadic Adjustment Scale" (DAS) und des "Partnerschaftsfragebogens" (PFB). [Assessment of relationship satisfaction. A comparison of the "Dyadic Adjustment Hahlweg, K. & Markman, H. J. (1988). Effectiveness of behavioral marital therapy: Empirical status of behavioral techniques in preventing and alleviating marital distress]. *Journal of Consulting and Clinical Psychology, 56,* 440–447.

Hahlweg, K., & Richter, D. (2010). Prevention of marital instability and distress. Results of an 11-year longitudinal follow-up study. *Behaviour Research and Therapy, 48,* 377–383. http://dx.doi.org/10.1016/j.brat.2009.12.010

Hammen, C. (1991). Generation of stress in the course of unipolar depression. *Journal of Abnormal Psychology, 100,* 555–561. http://dx.doi.org/10.1037/0021-843X.100.4.555

Hammen, C. (2006). Stress generation in depression: Reflections on origins, research, and future directions. *Journal of Clinical Psychology, 62,* 1065–1082. http://dx.doi.org/10.1002/jclp.20293

Hammen, C., & Brennan, P. A. (2002). Interpersonal dysfunction in depressed women: Impairments independent of depressive symptoms. *Journal of Affective Disorders, 72,* 145–156. http://dx.doi.org/10.1016/S0165-0327(01)00455-4

Hautzinger, M., & Bailer, M. (1992). *Allgemeine Depressionsskala: Manual* [Center for Epidemiological Studies Depression Scale]. Göttingen, Germany: Hogrefe.

Heyman, R. E. (2001). Observation of couple conflicts: Clinical assessment applications, stubborn truths, and shaky foundations. *Psychological Assessment, 13,* 5–35. http://dx.doi.org/10.1037/1040-3590.13.1.5

Heyman, R. E., Hunt-Martorano, A. N., Malik, J., & Slep, A. M. S. (2009). Desired change in couples: Gender differences and effects on communication. *Journal of Family Psychology, 23,* 474–484. http://dx.doi.org/10.1037/a0015980

Hothorn, T., Hornik, K., Strobl, C., & Zeileis, A. (2011). *A laboratory for recursive parti- tioning*. Retrieved from http://cran.r-project.org/web/packages/party/party.pdf

Hothorn, T., Hornik, K., & Zeileis, A. (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics, 15,* 651–674. http://dx.doi.org/10.1198/106186006X133933

Johannes, B., Wittels, P., Enne, R., Eisinger, G., Castro, C. A., Thomas, J. L., . . . Gerzer, R. (2007). Non-linear function model of voice pitch dependency on physical and mental load. *European Journal of Applied Physiology, 101,* 267–276. http://dx.doi.org/10.1007/s00421-007-0496-6

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin, 129,* 770–814.

Juslin, P., & Scherer, K. (2005). Vocal expression of affect. In J. Harrigan, R. Rosenthal, & K. R. Scherer (Eds.), *The new handbook of methods in nonverbal behavioral research* (pp. 65–136). New York, NY: Oxford University Press.

Kaiser, A., Hahlweg, K., Fehm-Wolfsdorf, G., & Groth, T. (1998). The efficacy of a compact psychoeducational group training program for married couples. *Journal of Consulting and Clinical Psychology, 66,* 753–760. http://dx.doi.org/10.1037/0022-006X.66.5.753

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association, 53,* 457–481. http://dx.doi.org/10.1080/01621459.1958.10501452

Karney, B. R., & Bradbury, T. N. (1995). The longitudinal course of marital quality and stability: A review of theory, method, and research. *Psychological Bulletin, 118,* 3–34. http://dx.doi.org/10.1037/0033-2909.118.1.3

Kiecolt-Glaser, J. K., Bane, C., Glaser, R., & Malarkey, W. B. (2003). Love, marriage, and divorce: Newlyweds' stress hormones foreshadow relationship changes. *Journal of Consulting and Clinical Psychology, 71,* 176–188. http://dx.doi.org/10.1037/0022-006X.71.1.176

Kiecolt-Glaser, J. K., & Newton, T. L. (2001). Marriage and health: His and hers. *Psychological Bulletin, 127,* 472–503. http://dx.doi.org/10.1037/0033-2909.127.4.472

Kirschbaum, C., & Hellhammer, D. H. (1994). Salivary cortisol in psychoneuroendocrine research: Recent developments and applications. *Psychoneuroendocrinology, 19,* 313–333. http://dx.doi.org/10.1016/0306-4530(94)90013-2

Kornhuber, J., Zenses, E. M., Lenz, B., Stoessel, C., Bouna-Pyrrou, P., Rehbein, F., . . . Mößle, T. (2013). Low 2D:4D values are associated with video game addiction. *PLoS ONE, 8*(11), e79539. http://dx.doi.org/10.1371/journal.pone.0079539

Kraemer, H. C., Morgan, G. A., Leech, N. L., Gliner, J. A., Vaske, J. J., & Harmon, R. J. (2003). Measures of clinical significance. *Journal of the American Academy of Child and Adolescent Psychiatry, 42,* 1524–1529. http://dx.doi.org/10.1097/00004583-200312000-00022

Larsen, J. T., Berntson, G. G., Poehlmann, K. M., Ito, T. A., & Cacioppo, J. T. (2008). The psychophysiology of emotion. In R. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *The handbook of emotions* (3rd ed., pp. 180–195). New York, NY: Guilford Press.

Margolin, G., Talovic, S., & Weinstein, C. D. (1983). Areas of Change Questionnaire: A practical approach to marital assessment. *Journal of Consulting and Clinical Psychology, 51,* 920–931. http://dx.doi.org/10.1037/0022-006X.51.6.920

Markman, H. J., Renick, M. J., Floyd, F. J., Stanley, S. M., & Clements, M. (1993). Preventing marital distress through communication and conflict management training: A 4- and 5-year follow-up. *Journal of Consulting and Clinical Psychology, 61,* 70–77. http://dx.doi.org/10.1037/0022-006X.61.1.70

Markman, H. J., Rhoades, G. K., Stanley, S. M., & Peterson, K. M. (2013). A randomized clinical trial of the effectiveness of premarital intervention: Moderators of divorce outcomes. *Journal of Family Psychology, 27,* 165–172. http://dx.doi.org/10.1037/a0031134

Markman, H. J., Rhoades, G. K., Stanley, S. M., Ragan, E. P., & Whitton, S. W. (2010). The premarital communication roots of marital distress and divorce: The first five years of marriage. *Journal of Family Psychology, 24,* 289–298. http://dx.doi.org/10.1037/a0019481

Markman, H. J., Stanley, S. M., & Blumberg, S. L. (2010). *Fighting for your marriage*. San Francisco, CA: Jossey-Bass.

Mauss, I. B., Levenson, R. W., McCarter, L., Wilhelm, F. H., & Gross, J. J. (2005). The tie that binds? Coherence among emotion experience, behavior, and physiology. *Emotion, 5,* 175–190. http://dx.doi.org/10.1037/1528-3542.5.2.175

McKnight, P. E., McKnight, K. M., Sidani, S., & Figueredo, A. J. (2007). *Missing Data: A gentle introduction*. New York, NY: Guilford Press.

Osborne, J. W., & Overbay, A. (2004). The power of outliers (and why researchers should always check for them). *Practical Assessment, Research & Evaluation, 9,* 1–10.

Owren, M. J., & Bachorowski, J. (2007). Measuring emotion-related vocal acoustics. In J. A. Coan & J. J. B. Allen (Eds.), *Handbook of emotion elicitation and assessment* (pp. 239–266). Oxford, UK: Oxford University Press.

Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1,* 385–401. http://dx.doi.org/10.1177/014662167700100306

Raftery, A. E. (1995). Bayesian model selection in social research (with discussion). In P. V. Marsden (Ed.), *Sociological methodology 1995* (pp. 111–195). Cambridge, MA: Blackwell Publishers.

R Development Core Team. (2012). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., & Müller, M. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics, 12,* 77. http://dx.doi.org/10.1186/1471-2105-12-77

Robles, T. F., & Kiecolt-Glaser, J. K. (2003). The physiology of marriage: Pathways to health. *Physiology & Behavior, 79,* 409–416. http://dx.doi.org/10.1016/S0031-9384(03)00160-4

Robles, T. F., Slatcher, R. B., Trombello, J. M., & McGinn, M. M. (2014). Marital quality and health: A meta-analytic review. *Psychological Bulletin, 140,* 140–187. http://dx.doi.org/10.1037/a0031859

Rufibach, K. (2012). A smooth ROC curve estimator based on log-concave density estimates. *The International Journal of Biostatistics, 8*(1). Advance online publication. http://dx.doi.org/10.1515/1557-4679.1378

Sbarra, D. A., Law, R. W., & Portley, R. M. (2011). Divorce and death: A meta-analysis and research agenda for clinical, social, and health psychology. *Perspectives on Psychological Science, 6,* 454–474. http://dx.doi.org/10.1177/1745691611414724

Schumacher, M., Holländer, N., Schwarzer, G., & Sauerbrei, W. (2001). Prognostic factor studies. In J. Crowley (Ed.), *Statistics in oncology* (pp. 321–378). New York, NY: Marcel Dekker.

Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *Journal of Marriage and the Family, 38,* 15–28. http://dx.doi.org/10.2307/350547

Stekhoven, D. J. (2012). *Nonparametric missing value imputation using Random Forest. R package version 1.3*. Retrieved from http://cran.r-project.org/web/packages/missForest/index.html

Stekhoven, D. J., & Bühlmann, P. (2012). MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics (Oxford, England), 28,* 112–118. http://dx.doi.org/10.1093/bioinformatics/btr597

Swets, J. A. (1986). Indices of discrimination or diagnostic accuracy: Their ROCs and implied models. *Psychological Bulletin, 99,* 100–117. http://dx.doi.org/10.1037/0033-2909.99.1.100

Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science, 240,* 1285–1293. http://dx.doi.org/10.1126/science.3287615

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, UK: Addison-Wesely.

U.S. National Center for Health Statistics. (2010). *National Vital Statistics Reports (NVSR), birth, marriages, divorces and deaths: Provisional data for 2009* (vol. *58,* no. 25). Retrieved from http://www.census.gov/compendia/statab/2012/tables/12s0133.pdf

van Buuren, S., & Groothuis-Oudshoorn, K. (2011a). *MICE: Multivariate imputation by chained equations. R package version 2.9*. Retrieved from http://CRAN.R- project.org/package=mice

van Buuren, S., & Groothuis-Oudshoorn, K. (2011b). MICE: Multivariate imputation by chained equations. *Journal of Statistical Software, 45,* 1–67.

Walfish, S. (2006). A review of statistical outlier methods. *Pharmaceutical Technology, 30,* 1–5.

Weitzman, E. D., Fukushima, D., Nogeire, C., Roffwarg, H., Gallagher, T. F., & Hellman, L. (1971). Twenty-four hour pattern of the episodic secretion of cortisol in normal subjects. *The Journal of Clinical Endocrinology and Metabolism, 33,* 14–22. http://dx.doi.org/10.1210/jcem-33-1-14

Weusthoff, S., Baucom, B. R., & Hahlweg, K. (2013a). Fundamental frequency during couple conflict: An analysis of physiological, behavioral, and sex-linked information encoded in vocal expression. *Journal of Family Psychology, 27,* 212–220. http://dx.doi.org/10.1037/a0031887

Whisman, M. (2001). The association between depression and marital distress. In S. Beach (Ed.), *Marital and family processes in depression: A scientific foundation for clinical practice* (pp. 3–24). Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/10350-001

Wilcox, R. (2012). *Modern statistics for the Social and Behavioral Sciences: A practical introduction*. New York, NY: Chapman & Hall/CRC press.

Woodin, E. M. (2011). A two-dimensional approach to relationship conflict: Meta-analytic findings. *Journal of Family Psychology, 25,* 325–335. http://dx.doi.org/10.1037/a0023791

Received May 15, 2014
Revision received June 16, 2015
Accepted June 16, 2015 ■