

Evidence-Based Approaches to Assessing Couple Distress

Douglas K. Snyder
Texas A&M University

Richard E. Heyman
Stony Brook University, State University of New York

Stephen N. Haynes
University of Hawai'i at Manoa

This article describes a conceptual framework for couple-based assessment strategies grounded in empirical findings linking couple distress to a broad range of both individual and relationship characteristics. These characteristics can contribute to, exacerbate, or result from relationship problems. On the basis of these findings, the authors articulate specific targets of clinical inquiry reflecting relationship behaviors, cognitions, and affect as well as features of individual distress. Guided by this framework, empirically supported assessment strategies and techniques emphasizing relationship functioning across diverse methods are proposed, including the clinical interview, analog behavioral observation, and both self- and other-report measures. Discussion concludes with specific recommendations regarding clinical assessment of couple distress and directions for further research.

Keywords: couples, marriage, relationships, assessment, behavioral observation

Couple distress, or “partner relational problems,” as defined by the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; *DSM-IV*; American Psychiatric Association, 1994, p. 681), involves communication or related interactional difficulties associated with clinically significant impairment in individual or relationship functioning or the development of symptoms in one or both partners. Couple assessment is distinguished from individual assessment by its emphasis on relationship processes that interact with but transcend individual characteristics. Although most individual assessment strategies offer at least token acknowledgment of the potential contributing role of relationship and broader socioecological factors, they do not emphasize relationship issues. By contrast, in couple-based assessment, relationship processes are the initial focal point of conceptualization and measurement strategies, with individual and extended social system processes integrated as critical but second-order components.

Couple distress is prevalent in both community epidemiological studies and in research involving individual treatment samples. In the United States, the most salient indicator of couple distress remains a divorce rate of approximately 50% among married

couples (Kreider & Fields, 2002), with about half of these occurring within the first 7–8 years of marriage. Independent of divorce, the research literature suggests that many marriages, if not most, experience periods of significant turmoil that place them at risk for dissolution or symptom development (e.g., depression or anxiety) in one or both partners at some point in their lives. Only one third of married persons report being “very happy” with their marriage, down from more than one half 25 years ago. About 37% of men aged 50 to 59 and 20% of women aged 40 to 49 report having had an affair at least once during their marriage (Laumann, Gagnon, Michael, & Michaels, 1994). An early study indicated that more people sought therapy for marital problems than for any other type of problem (Veroff, Kulka, & Douvan, 1981). In a recent national survey, the most frequently cited causes of acute emotional distress were relationship problems, including divorce, separation, and other marital strains (Swindle, Heller, Pescosolido, & Kikuzawa, 2000). Other recent studies indicate that maritally discordant individuals are overrepresented among individuals seeking mental health services, regardless of whether they report marital distress as their primary complaint (Lin, Goering, Offord, Campbell, & Boyle, 1996). In a study of 800 employee assistance program clients, 65% rated family problems as “considerable” or “extreme” (Shumway, Wampler, Dersch, & Arredondo, 2004). Moreover, a study by Law and Crane (2000) of medical records for 292 patients from a health-maintenance organization showed that those patients who received marital or family therapy significantly reduced their use of health care services by 21.5%, indicating a substantial cost-offset effect for such therapy.

The linkage of relationship distress to disruption of individual emotional and physical well-being further emphasizes the importance of integrating empirically based strategies for assessing couple distress as an adjunct to effective intervention strategies with individuals. Research clearly indicates that couple distress covaries with individual emotional and behavioral disorders above

Douglas K. Snyder, Department of Psychology, Texas A&M University; Richard E. Heyman, Department of Psychology, Stony Brook University, State University of New York; Stephen N. Haynes, Department of Psychology, University of Hawai'i at Manoa.

Richard E. Heyman's contribution to this article was supported by Grant R01-MH06704301, from the National Institute of Mental Health.

Correspondence concerning this article should be addressed to Douglas K. Snyder, Department of Psychology, Texas A&M University, 4235 TAMU, College Station, TX 77843-4235; Richard E. Heyman, Department of Psychology, Stony Brook University, State University of New York, Stony Brook, NY 11794-2500; or Stephen N. Haynes, Department of Psychology, University of Hawai'i at Manoa, 2430 Campus Road, Honolulu, HI 96822. E-mail: d-snyder@tamu.edu, Richard.Heyman@Stonybrook.edu, or sneil@hawaii.edu

and beyond general distress in other close relationships (Whisman, Sheldon, & Goering, 2000). Moreover, couple distress—particularly negative communication—has direct adverse effects on cardiovascular, endocrine, immune, neurosensory, and other physiological systems that, in turn, contribute to physical health problems (Kiecolt-Glaser & Newton, 2001). Also, the effects of couple distress are not confined to the adult partners. Gottman (1999) cites evidence indicating that “marital distress, conflict, and disruption are associated with a wide range of deleterious effects on children, including depression, withdrawal, poor social competence, health problems, poor academic performance, a variety of conduct-related difficulties, and markedly decreased longevity” (p. 4). In brief, couple distress has a markedly high prevalence; has a strong linkage to emotional, behavioral, and health problems in the adult partners and their offspring; and is among the most frequent primary or secondary concerns reported by individuals seeking assistance from mental health professionals.

In this article, we first propose a conceptual framework for couple-based assessment strategies grounded in empirical findings linking couple distress to a broad range of both individual and relationship characteristics that have been shown to contribute to, covary with, or result from relationship problems. Guided by this conceptual framework, we then describe specific, empirically supported assessment strategies and techniques emphasizing relationship functioning across diverse methods including interview, observation, and both self- and other-report. On the basis of previous findings, we argue that attention should be given to disorders of individual emotional or behavioral functioning when clinicians assess couple distress and to relationship distress when they treat individual disorders. Finally, we conclude with specific recommendations regarding assessment of couple distress and directions for further research.

Conceptualizing Couple Distress: Empirical Foundations

Previous research has shown couple distress to relate to a broad range of individual and relationship characteristics. Efficacious assessment of these characteristics requires a conceptual framework for directing both assessment domains and methods. Snyder, Cavell, Heffer, and Mangrum (1995) proposed a multitrait, multilevel assessment model composed of five overlapping construct domains (cognitive, affective, behavioral, interpersonal, and structural/developmental) operating at five system levels (individuals, dyads, the nuclear family, the extended family, and community/cultural systems). Table 1 (from Snyder & Abbott, 2002) provides a modest sampling of specific constructs relevant to each domain at each system level.

The relevance of any specific facet of the model to relationship distress for either partner varies dramatically across couples; hence, although Table 1 provides guidance regarding initial areas of inquiry from a nomothetic perspective, the relation of any specific component to relationship distress for a given individual or couple needs to be determined from a functional analytic approach and applied idiographically (Cone, 1988; Haynes, Leisen, & Blaine, 1997; Haynes & O'Brien, 2000). Moreover, interactive effects occur within domains across levels, within levels across domains, and across levels and domains. For example, partners' expectancies regarding respective parenting roles likely relate in part to both their respective family models and broader

cultural norms; moreover, lack of congruence in such expectancies may contribute to relationship conflict and various coercive means of influencing each other's behavior if communication skills for articulating relationship concerns and negotiating satisfactory resolution are deficient.

It lies outside the scope of this article to review all the correlates of couple distress within each domain at each system level and the findings regarding their interaction. Instead, we will highlight more salient components operating primarily at the dyadic level. An extended discussion of individual and relationship characteristics as they relate to both assessment and treatment of couple distress can be found in Epstein and Baucom (2002).

Relationship Behaviors

Research examining behavioral components of couple distress has emphasized two domains: the rates and reciprocity of positive and negative behaviors exchanged between partners and communication behaviors related to both emotional expression and decision making. Regarding the former, distressed couples are distinguished from nondistressed couples by multiple characteristics, including (a) higher rates of negative verbal and nonverbal exchanges (e.g., disagreements, criticism, hostility); (b) higher levels of reciprocity in negative behavior (i.e., the tendency for negativity in Partner A to be followed by negativity in Partner B); (c) lengthier chains of negative behavior once initiated; (d) higher ratios of negative to positive behaviors, independent of their separate rates; and (e) lower rates of positive verbal and nonverbal behaviors, for example, approval, empathy, smiling, positive touch (Weiss & Heyman, 1997). Findings suggest a stronger linkage for negativity, compared with positivity, to overall couple distress.

Given the inevitability of disagreements arising in long-term relationships, numerous studies have focused on specific communication behaviors that exacerbate or impede the resolution of couple conflicts. Most notable among these are difficulties in articulating thoughts and feelings related to specific relationship concerns and deficits in decision-making strategies for containing, reducing, or eliminating conflict. Gottman (1994) observed that specific communication behaviors involving expression of criticism and contempt, along with defensiveness and withdrawal, predicted long-term distress and risk for relationship dissolution. Christensen and Heavey (1990) found that distressed couples were more likely than nondistressed couples to demonstrate a demand → withdraw pattern, in which one person attempts to engage the partner in relationship exchange and that partner withdraws, with respective approach and retreat behaviors progressively intensifying.

Given findings regarding the prominence of negativity, conflict, and ineffective decision-making strategies as correlates of relationship distress, couple assessment must address specific questions regarding relationship behaviors. We list these questions below, along with sample assessment methods; in subsequent sections specifying interview, observational, and self-report strategies for assessing couple distress, we describe these and related methods in greater detail.

1. How frequent and intense are the couple's conflicts? How rapidly do initial disagreements escalate into major arguments? For how long do conflicts persist without resolution? Both inter-

Table 1
Sample Assessment Constructs Across Domains and Levels of Couple and Family Functioning

Domain	Individual	Dyad (couple, parent-child)	Nuclear family system	Extended system (family of origin, friends)	Culture/community
Cognitive	Intelligence; memory functions; thought content; thought quality; analytic skills; cognitive distortions; schemas; capacity for self-reflection and insight.	Cognitions regarding self and other in relationship; expectancies, attributions, attentional biases, and goals in the relationship.	Shared or co-constructed meanings within the system; family ideology or paradigm; thought sequences between members contributing to family functioning.	Intergenerational patterns of thinking and believing; co-constructed meaning shared by therapist and family or other significant friends or family.	Prevailing societal and cultural beliefs and attitudes; ways of thinking associated with particular religious or ethnic groups that are germane to the family or individual.
Affective	Mood; affective range, intensity, and valence; emotional lability.	Predominant emotional themes or patterns in the relationship; cohesion; range of emotional expression; commitment and satisfaction in the relationship; emotional content during conflict; acceptance and forgiveness.	Family emotional themes of fear, shame, guilt, or rejection; system properties of cohesion or emotional disaffection; emotional atmosphere in the home - including humor and joy as well as conflict and hostility.	Emotional themes and patterns in extended system; intergenerational emotional legacies; patterns of fusion or differentiation across generations.	Prevailing emotional sentiment in the community, culture, and society; cultural norms and mores regarding the expression of emotion.
Behavioral	Capacity for self-control; impulsivity; aggressiveness; capacity to defer gratification. Overall energy and drive.	Recursive behavioral sequences displayed in the relationship; behavioral repertoire; reinforcement contingencies; strategies used to control other's behavior.	Repetitive behavioral patterns or sequences used to influence family structure and power; shared recreation and other pleasant activities.	Behavioral patterns displayed by the extended system (significant friends, family of origin, therapist) used to influence the structure and behaviors of the extended system.	Cultural norms and mores of behavior; behaviors which are prescribed or proscribed by the larger society.
Interpersonal/communication	Characteristic ways of communicating and interacting across relationships or personality (e.g. shy, gregarious, narcissistic, dependent, controlling, avoidant).	Quality and frequency of the dyad's communication; speaking and listening skills; how couples share information, express feelings, and resolve conflict.	Information flow in the family system; paradoxical messages; family system boundaries, hierarchy, and organization; how the family system uses information regarding its own functioning; family decision-making strategies.	Degree to which information is shared with and received from significant others outside the nuclear family system or dyad; has implications about the relative permeability of boundaries and the degree to which the family or couple is receptive to outside influences.	Information that is communicated to the family or individual by the community or culture in which they live; how the family or individual communicates their needs and mobilizes resources.
Structural/developmental	All aspects of physiological and psychosocial development; personal history that influences current functioning; intrapersonal consistency of cognitions, affect, and behavior.	History of the relationship and how it has evolved over time; congruence of partners' cognitions, affect, and behavior.	Changes in the family system over time; current stage in the family life cycle; congruence in needs, beliefs, and behaviors across family members.	Developmental changes across generations; significant historical events influencing current system functioning (e.g. death, illness, divorce, abuse); congruence of beliefs and values across extended social support systems.	The cultural and political history of the society in which the family or individual lives; current political and economic changes; congruence of the individual's or couple's values with those of the larger community.

Note. From *Handbook of Assessment and Treatment Planning for Psychological Disorders* by M. M. Antony and D. H. Barlow (Eds.), 2002, New York: Guilford Press. Copyright 2002 by Guilford Press. Reprinted with permission.

view and self-report measures may yield useful information regarding rates and intensity of negative exchanges as well as patterns of conflict engagement. Commonly used self-report measures specific to communication include the Communication Patterns Questionnaire (CPQ; Christensen, 1987) and Styles of Conflict Inventory (SCI; Metz, 1993). Couples' conflict-resolution patterns may be observed directly by instructing partners to discuss problems of their own choosing representative of both moderate and high disagreement and then either formally or informally coding these interactions using one of the behavioral coding systems described later in this article.

2. What are common sources of relationship conflict? Examples may include interactions regarding finances, children, sexual intimacy, use of leisure time, or household tasks; involvement with others including extended family, friends, or coworkers; and differences in preferences or core values. In addition to the clinical interview, numerous self-report measures sample sources of distress across a variety of relationship domains; among these are the Areas of Change Questionnaire (ACQ; Weiss & Birchler, 1975), Areas of Change Checklist (Gottman, 1999), Marital Satisfaction Inventory—Revised (MSI-R; Snyder, 1997), and Spouse Observation Checklist (SOC; Birchler, Weiss, & Vincent, 1975).

3. What resources and deficits do partners demonstrate in problem-identification and conflict-resolution strategies? Do they engage couple issues at adaptive levels—that is, neither avoiding nor dwelling on relationship concerns? Do partners balance their expression of feelings with decision-making strategies? Are problem-resolution efforts hindered by inflexibility or imbalances in power? Do partners offer each other support when confronting stressors from within or outside their relationship? As noted by others (e.g., Bradbury, Rogge, & Lawrence, 2001; Cutrona, 1996), most of the interactional tasks developed for use in couple research have emphasized problem solving and conflict resolution to the exclusion of tasks designed to elicit more positive relationship behaviors such as emotional or strategic support. Hence, when designing interaction tasks for couples, both clinicians and researchers should include tasks specifically designed to sample potential positive, as well as negative, exchanges. For example, couples might be asked to discuss a time when one partner's feelings were hurt by someone outside the relationship (e.g., a friend or coworker) in order to assess behaviors expressing understanding and caring.

Relationship Cognitions

Social learning models of couple distress have expanded to emphasize the role of cognitive processes in moderating the impact of specific behaviors on relationship functioning (Baucom, Epstein, & LaTaillade, 2002). Research in this domain has focused on such factors as selective attention, attributions for positive and negative relationship events, and specific relationship assumptions, standards, and expectancies. For example, findings indicate that distressed couples often exhibit a bias toward selectively attending to negative partner behaviors and relationship events and ignoring or minimizing positive events (Sillars, Roberts, Leonard, & Dun, 2000). Compared with nondistressed couples, distressed partners also tend to blame each other for problems and to attribute each other's negative behaviors to broad and stable traits (Bradbury & Fincham, 1990). Distressed couples are also more likely to

have unrealistic standards and assumptions about how relationships should work and lower expectancies regarding their partner's willingness or ability to change their behavior in some desired manner (Epstein & Baucom, 2002). On the basis of these findings, assessment of relationship cognitions should emphasize the following questions:

1. Do partners demonstrate an ability to accurately observe and report both positive and negative relationship events? For example, partners' descriptions and interpretations of couple interactions observed directly in therapy can be compared with the clinician's own assessment of these same exchanges. Partners' response sets from completed self-report relationship measures can also be assessed; for example, the Conventionalization (CNV) scale on the MSI-R assesses the tendency to distort relationship appraisals in an overly positive direction.

2. What interpretation or meaning do partners impart to relationship events? The clinical interview is particularly useful for eliciting partners' subjective interpretations of their own and each other's behaviors; such interpretations and attributions also frequently are expressed during conflict-resolution or other interactional tasks. To what extent are partners' negative relationship behaviors attributed to stable, negative aspects of the partner versus to external or transient events? Self-report measures assessing relationship attributions include the Dyadic Attributional Inventory (DAI; Baucom, Sayers, & Duhe, 1989), the Marital Attitude Survey (Pretzer, Epstein, & Fleming, 1992), and the Relationship Attribution Measure (Fincham & Bradbury, 1992).

3. What beliefs and expectancies do partners hold regarding both their own and the other person's ability and willingness to change in a manner anticipated to be helpful to their relationship? What standards do they hold for relationships generally? In addition to assessing such cognitions using the clinical interview, clinicians can assess these and similar relationship expectancies with such self-report measures as the Inventory of Specific Relationship Standards (ISRS; Baucom, Epstein, Rankin, & Burnett, 1996) and the Relationship Beliefs Inventory (Eidelson & Epstein, 1982).

Relationship Affect

Similar to findings regarding behavior exchange, research indicates that distressed couples are distinguished from nondistressed couples by higher overall rates, duration, and reciprocity of negative relationship affect and, to a lesser extent, by lower rates of positive relationship affect. Nondistressed couples show less reciprocity of positive affect, reflecting partners' willingness or ability to express positive sentiment spontaneously independent of their partner's affect (Gottman, 1999). By contrast, partners' influence on each other's negative affect has been reported for both proximal and distal outcomes. For example, Pasch, Bradbury, and Davila (1997) found that partners' negative mood prior to discussion of a personal issue predicted lower levels of emotional support they provided to the other during their exchange. From a longitudinal perspective, couples who divorce are distinguished from those who remain married by partners' initial levels of negative affect and by a stronger linkage of initial negativity to the other person's negative affect over time (Cook et al., 1995). Gottman (1999) determined that the single best predictor of couples' eventual divorce was the amount of contempt partners ex-

pressed in videotaped interactions. Hence, assessment of couple distress should determine the following:

1. To what extent do partners express and reciprocate negative and positive feelings about their relationship and toward each other? Partners' reciprocity of affect is best observed directly during either structured or unstructured interactions. Separate from observational strategies are numerous self-report measures tapping such affective components of relationship functioning as satisfaction (e.g., the Dyadic Adjustment Scale; DAS; Spanier, 1976); the MSI-R; and the Relationship Satisfaction Scale, Burns & Sayers, 1992) and intimacy (e.g., the Personal Assessment of Intimacy in Relationships; Schaefer & Olson, 1981). Although much of the couple literature emphasizes negative emotions, it is equally important for positive emotions, such as smiling, laughter, expressions of appreciation or respect, comfort or soothing, and similar expressions, to be assessed through observation or clinical inquiry.

2. What ability does each partner have to express his or her feelings in a modulated manner? Problems with emotion self-regulation may be observed either in overcontrol of emotions (e.g., an inability to access, label, or express either positive or negative feelings) or in undercontrol of emotions (e.g., the rapid escalation of anger into intense negativity approaching rage, progression of tearfulness into sobbing, or deterioration in quality of thought secondary to emotional overload). Self-report measures assessing emotion regulation in relationships include the Managing Affect and Differences Scale (Arellano & Markman, 1995) and a recently developed Difficulties in Emotion Regulation Scale (Gratz & Roemer, 2004).

3. To what extent does partners' negative affect generalize across occasions? Generalization of negative affect, or "negative sentiment override" (Weiss, 1980), can be observed in partners' inability to shift from negative to either neutral or positive affect during the interview or interactional tasks or in reports of distress across most or all domains of relationship functioning assessed using self-report. In research applications, ratings of affect by partners observing their videotaped interactions may provide an additional means of assessing sentiment override. For example, in a study of the effects of relationship sentiment override on couples' perceptions, partners used an affect-rating dial to indicate how positively or negatively they felt during a previously videotaped interaction and how they thought their partner felt during the interaction (Hawkins, Carrère, & Gottman, 2002).

Individual Distress

Separate from these relationship processes characterizing couple distress, there is growing evidence that relationship difficulties covary with, contribute to, and result from individual emotional and behavioral disorders (Snyder & Whisman, 2003). Whisman (1999) and Whisman and Uebelacker (2003) found that maritally distressed people are more likely to have psychiatric disorders than are nondistressed people, that this association extends across diverse emotional and behavioral disorders, and that the magnitudes of these associations are generally quite large. Similarly, Whisman et al. (2000) determined that marital distress was associated with six specific disorders (i.e., major depression, social and simple phobia, panic disorder, generalized anxiety disorder, and alcohol dependence or abuse) above and beyond general distress in other close relationships. Although findings linking marital distress to

individual differences in emotional or behavioral functioning within the nondistressed range are less strong, Bradbury and Karney (1993) determined that neuroticism—that is, a general tendency to experience the world negatively—was associated with marital distress.

Although anecdotal evidence from therapist surveys (Northey, 2002; Whisman, Dixon, & Johnson, 1997) suggests that individual difficulties render couple therapy more difficult or less effective, empirical findings supporting this have been limited, in part because of exclusionary criteria frequently implemented in controlled treatment outcome studies. An exception involves studies showing that depression in one or both spouses predicts poorer response to couple therapy (Sher, Baucom, & Larus, 1990; Snyder, Mangrum, & Wills, 1993). Moreover, a presenting problem relating to an individual dysfunction predicts premature dropout from couple therapy (Allgood & Crane, 1991).

The evidence regarding the negative impact of couple distress on the treatment of individual emotional or behavioral problems is somewhat stronger, affirming the importance of clinicians specifically assessing couple distress when treating individual psychopathology. For example, among married people treated for generalized anxiety disorder, the likelihood of reduction in anxiety symptoms decreases as marital distress increases (Durham, Allan, & Hackett, 1997). Similarly, studies of treatment for agoraphobia indicate that higher relationship distress at pretreatment predicts a poorer response to treatment (Daiuto, Baucom, Epstein, & Dutton, 1998). Several studies have shown that marital distress is associated with slower recovery in treatment for depression (Goering, Lancee, & Freeman, 1992) and a greater likelihood of relapse (e.g., Hooley & Teasdale, 1989; Whisman, 2001). Finally, several studies have found couple relationship problems to predict poorer response to alcohol- and drug-abuse treatment programs (see Fals-Stewart, Birchler, & O'Farrell, 2003, for a review). On the basis of these findings, attention should be given to disorders of individual emotional or behavioral functioning when clinicians assess couple distress and to relationship distress when they treat individual disorders.

1. To what extent does either partner exhibit individual emotional or behavioral difficulties potentially contributing to, exacerbating, or resulting in part from couple distress? Given the association of couple distress with affective disorders and alcohol use, initial interviews of couples should include questions regarding suicidality and alcohol or other substance use as well as brief screening for previous treatment of emotional or behavioral disorders. Because individuals entering couple therapy are often apprehensive regarding inferences concerning potential individual psychopathology, such measures as the Minnesota Multiphasic Personality Inventory (Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989) or Personality Assessment Inventory (Morey, 1991) may provoke defensiveness or disrupt initial efforts to establish a collaborative therapeutic alliance. Hence, when a clinical interview suggests potential interaction of relationship and individual dysfunction, more focused and brief measures (e.g., the Beck Depression Inventory—II; Beck, Steer, & Brown, 1996, or the Symptom Checklist-90-Revised; Derogatis & Savitz, 1999) should be considered.

2. In formulating treatment strategies, to what extent would either partner benefit from interventions targeting individual difficulties either within couple-based strategies or in separate indi-

vidual treatment? Guidelines for recommending collateral treatment of individual disorders when approaching conjoint treatment of couple distress lie outside the scope of this article. Considerations include the nature and extent of individual dysfunction, anticipated impact of individual difficulties on couple treatment, the availability of effective interventions (individual, conjoint, or pharmacological) for the disorder in question, and the receptiveness of either partner to such recommendations. Discussions of such factors have been presented elsewhere (Baucom, Shoham, Mueser, Daiuto, & Stickle, 1998; Snyder & Whisman, 2003).

3. Given findings regarding the impact of couple distress on treatment of individual disorders, as well as the favorable impact of couple and family therapy on utilization of health services (Law & Crane, 2000), screening for relationship distress should be routine when individuals are evaluated for emotional, behavioral, or physical health problems. Given the constraints of such assessment in routine practice, we advocate a sequential strategy of progressively more detailed assessment when indicators of relationship distress emerge (cf., Snyder & Abbott, 2002, pp. 366–367):

- (a) Conduct clinical inquiry as to whether relationship problems contribute to the individual feeling depressed or anxious or feeling less able to deal with such stresses as work, children and family, or health concerns.
- (b) Alternatively, use an initial brief screening measure (e.g., the Kansas Marital Satisfaction Scale [KMSS]; Schumm et al., 1986; or Quality of Marriage Index [QMI]; Norton, 1983) having evidence of both reliability and discriminative validity.
- (c) For individuals reporting moderate to high levels of global relationship distress, follow up with a multidimensional self-report measure (e.g., the MSI-R) to differentiate among levels and sources of distress.

When screening for either clinical or research purposes, we advocate assessment strategies favoring sensitivity over specificity to minimize the likelihood of overlooking potential factors contributing to individual or relationship distress. This implies initial use of broad screening items in clinical inquiry or self-report measures, subsequent use of more extensive narrow-band measures to pinpoint specific sources of concern, followed by functional-analytic assessment strategies to delineate the manner in which individual and relationship concerns affect each other and relate to situational factors.

Additional Factors in Assessing Couple Distress

Separate from these domains of individual and relationship functioning or broader aspects of couples' socioecological system suggested by Table 1, several additional factors in assessing couple distress warrant consideration. First, as with individual disorders, assessment of distressed couples needs to be tailored to the specific constellation of partners' presenting complaints and the intrapersonal and relational contexts in which these are embedded as well as to the theoretical framework from which prevention or treatment interventions will likely proceed. Second, although we have emphasized primarily individual and relational factors contributing

to couple distress, it is equally important to assess couples' strengths and resources across intrapersonal, relationship, and broader social system levels. These include partners' ability to limit the impact of individual or couple dysfunction despite overwhelming stressors or to contain the generalization of distress to other family members. Third, couple assessment is distinguished from individual assessment in part by partners' tendency to attribute blame for relationship problems to each other rather than to themselves. In turn, partners' emotional reactivity can render valid assessment more difficult. Hence, principles of therapeutic assessment (Finn & Tonsager, 1997) emphasizing collaborative definition of assessment goals may be particularly helpful when the clinician is facilitating partners' goals for themselves and their relationship through relevant assessment strategies.

Finally, although explicitly noted by our conceptual framework, we wish to emphasize the importance of attending to cultural differences in the subjective experience and overt expression of couple distress as well as factors bearing on its development and treatment. By this, we refer not only to cross-national differences in couples' relationships but also to cross-cultural differences within nationality and consideration of nontraditional relationships, including gay and lesbian couples. There are important differences among couples as a function of their culture, religious orientation, economic level, and age. These dimensions can affect the importance of the couple relationship to a partner's quality of life, their expectancies regarding marital and parenting roles, typical patterns of verbal and nonverbal communication and decision making within the family, the behaviors that are considered distressing, sources of relationship conflict, the type of external stressors faced by a family, and the ways that partners respond to couple distress and divorce (e.g., Diener, Gohm, Suh, & Oishi, 2000; Gohm, Oishi, Darlington, & Diener, 1998; Jones & Chao, 1997). If we assume that the methods and focus of couples assessment are guided by assumptions about the factors that affect relationship satisfaction, it is likely that the optimal content validity of an instrument (i.e., the ideal set of items of a questionnaire or interview or the ideal observation scenarios or behavior codes in analog observation) would vary as a function of these dimensions. For example Haynes et al. (1992) found that parenting, extended family, and sex were less strongly related to marital satisfaction, whereas health of the spouse and other forms of affection were more important factors in marital satisfaction in older (i.e., over 55 years) couples compared with younger couples. Similarly, Bhugra and De Silva (2000) suggested that relationships with extended family members might be more important in some cultures. Also, when partners are from different cultures, cultural differences and conflicts can be a source of relationship dissatisfaction (e.g., Baltas & Steptoe, 2000).

In these examples, an instrument to measure relationship satisfaction, if developed for use with younger and European American couples, would likely show a satisfactory but less-than-optimal degree of validity and clinical utility when used with older persons or persons from non-European American cultures. If used with persons who differed from those used for instrument development, many items could be relevant, but some would likely not be relevant, and some important areas may not be tapped by the item pool. The ultimate consequences of using assessment instruments that are not appropriate for a client may include invalid clinical case formulations, decreased adherence to treatment recommenda-

tions, premature termination of therapy, and reduced treatment outcome (Tanaka-Matsumi, 2004).

As with individual disorders, assessment strategies for evaluating relationships vary across the clinical interview, observational methods, and self- and other-report measures. In the sections that follow, we discuss empirically supported techniques within each of these assessment strategies. Although specific techniques within any method could target diverse facets of individual, dyadic, or broader system functioning, we emphasize those more commonly used when couple distress is being assessed.

Interview-Based Methods

The clinical interview is the first step in assessing couples.¹ It can aid in identifying a couple's behavior problems and strengths, help specify a couple's treatment goals, and be used to acquire data that are useful for treatment outcome evaluation. The assessment interview can also serve to strengthen the client-assessor relationship, identify barriers to treatment, and increase the chance that the couple will participate in subsequent assessment and treatment tasks. Furthermore, it is the primary means of gaining a couple's informed consent about the assessment-treatment process. Data from initial assessment interviews also guide the assessor's decisions about which additional assessment strategies may be most useful; for example, Gordis, Margolin, and John (2001) used an interview to select topics for discussion during an analog behavioral observation (ABO) of couple communication patterns. Perhaps most important, the assessment interview can provide a rich source of hypotheses about factors that may contribute to the couple's distress. These hypotheses contribute to the case formulation, which, in turn, affects decisions about the best treatment strategy for a particular couple.

The interview can also be used to gather information on multiple levels, in multiple domains, and across multiple response modes in couple assessment. It can provide information on the specific behavioral interactions of the couple, including behavioral exchanges and violence; problem-solving skills; sources of disagreement; areas of satisfaction and dissatisfaction; each partner's thoughts, beliefs, and attitudes; and their feelings and emotions regarding the partner and relationship. The couple assessment interview also can provide information on cultural and family system factors and other events that might affect the couple's functioning and response to treatment. These factors might include interactions with extended family members, other relationship problems within the nuclear family (e.g., between parents and children), economic stressors, and health challenges. Finally, the initial assessment interview can provide information on potentially important causal variables for couple distress at an individual level, such as a partner's substance use, mood disorder, or problematic personality traits.

Given the central role of the initial assessment interview in couple therapy, the numerous outlines of recommended interview structures (e.g., Gottman, 1999; Karpel, 1994; L'Abate, 1994; Snyder & Abbott, 2002), and the historical advocacy for a more scholarly and empirical approach to assessment interviews (e.g., Coyne, 1986; Haynes & Jensen, 1979), it is surprising that couple assessment interviews have not been subjected to the rigorous development and psychometric evaluation that have characterized other couple assessment methods such as self-report questionnaires

(Snyder & Abbott, 2002), partner-report questionnaires (Weiss & Perry, 1983), and ABO (Heyman, 2001). Although elements of couple interviews (i.e., individual questions) have been developed and evaluated, comprehensive structured interviews for couple assessment have not been subjected to comprehensive psychometric evaluation.

Given the state of research regarding interviews for assessing couple distress, the goal of this section is to provide an overview of research on specific couple interview elements and a structure for developing and evaluating more comprehensive couple interview methods.

Domains of Interest, Content Validity, and Clinical Utility

Content validity refers to the degree to which elements of an assessment instrument are representative of, and relevant to, the targeted constructs (Haynes, Richard, & Kubany, 1995). In the case of couple interviews, content validity refers to the degree to which elements of a structured interview (e.g., individual questions, instructions, response formats) cover and are appropriate to the specific domains of interest and goals of the interview. Content validity can be more challenging for interviews than for other assessment methods because, as noted earlier, clinical interviews often have more goals and address more domains than do ABO or questionnaires.

There are many domains of interest in assessing couple distress, and an interview can have satisfactory content validity in one but not another domain. Moreover, each domain may have many facets, and each facet should ideally have several associated interview items or queries to enhance its measurement qualities. Although the interview is useful for describing couples in these domains, it can be especially useful in identifying functional relations that may account for relationship difficulties. The functional relations of greatest interest in couple assessment are those that are relevant to problem behaviors, feelings, and relationship enhancement. Identifying functional relations allows the assessor to hypothesize about why a partner is unhappy or what behavioral sequences lead to angry exchanges. Assessors are interested, for example, in finding out what triggers a couple's arguments and what communication patterns lead to their escalation. What does one partner do, or not do, that leads the other partner to feel unappreciated or angry?

Procedures for enhancing the content validity of assessment instruments early in their development have been outlined elsewhere in greater detail (Haynes et al., 1995; Haynes & O'Brien, 2000). The development of a structured couple interview would include several steps: (a) the preliminary specification of domains and facets of relationship distress; (b) attention to variables that have been found in research to be significantly related to relationship distress; (c) consultation with couples (e.g., focus groups) about the characteristics of their relationship and variables that affect it; (d) selection and refinement of items, on the basis of these multiple sources of input, to ensure adequate coverage of domains and facets; (e) review of early versions of the interview by experts;

¹ In this article, the *assessment interview* refers to an initial pretreatment interview rather than to treatment-outcome-related, feedback, debriefing, motivational, or intervention-oriented interviews.

and (f) preliminary indices of item performance (e.g., exploratory factor analyses or item response characteristics).

Such comprehensive interviews have been proposed (see Gottman's, 1999, 2-hr intake interview). For example, Heyman, Feldbau-Kohn, Ehrensaft, Langhinrichsen-Rohling, and O'Leary (2001) developed a structured interview to provide a diagnostic measure of relationship distress, aggression, and abuse. However, it would be challenging to develop a couple interview with satisfactory content validity for all of the relevant domains summarized in Table 1. A more feasible development process might involve a sequential approach, developing and validating a limited number of domains at a time. The multiple goals of the couple interview further complicate the development process. It would be especially important, for example, to establish the degree to which the interview increased a couple's participation in the intervention process and aided decisions about additional assessment strategies and case formulation.

The Validity of Partners' Self-Reports During a Couple Interview

Earlier we noted that no comprehensive structured couple interview has been subjected to formal psychometric evaluation. However, there are data on the predictive and concurrent validity of some partner reports from couple interviews. Several studies have found that spouse reports of marital satisfaction were significantly correlated with concurrent or future measures of psychological disorders. For example, Whisman (1999) reported that two interview items on marital satisfaction (rated on a 4-point scale) were significantly correlated with concurrent measures of major depression and posttraumatic stress disorder for women and dysthymia for men. The same items were also found to relate significantly to multiple indices of positive and negative marital interactions in another study (Forthofer, Markman, Cox, Stanley, & Kessler, 1996). Similarly, Whisman and Bruce (1999) found that in a sample of 904 community participants, an interview measure of marital satisfaction (based on a single item using a 3-point scale) was significantly correlated with the onset of major depressive disorder within the next year. Dissatisfied spouses, compared with satisfied spouses, were three times more likely to develop a major depressive disorder within the year following the initial interview. This single item was also significantly correlated with marital status (e.g., marital separation or divorce) 1 year later.

Leonard and Senchak (1996) found that an aggregate of four premarital interview items regarding husband aggression (verbal and physical aggression items taken from the Conflict Tactics Scale (CTS; Straus, 1979) was significantly correlated with measures of husband aggression obtained 1 year later. Haynes, Jensen, Wise, and Sherman (1981) found that initial intake interview items on communication and relationship satisfaction related significantly to questionnaire and ABO measures of similar constructs. Convergent validity indices were higher when participants were interviewed separately than when they were interviewed with their partner, particularly for sensitive items (e.g., those concerning sexual interactions).

Heyman et al. (2001) developed a structured interview patterned after the Structured Clinical Interview for the DSM (First, Gibbon, Spitzer, & Williams, 1997) to measure and diagnose relationship

distress, aggression, and abuse. Partners were interviewed separately. Findings from the interview significantly differentiated maritally distressed and happy couples, as measured by standardized self-report questionnaires (Marital Adjustment Test; Locke & Wallace, 1959). Specifically, couples diagnosed as maritally distressed or nondistressed also differed significantly on several codes derived from ABO discussions of problem topics. Heyman et al. (2001) also found a high level of concordance between CTS items presented in an interview and the same items given in the form of a questionnaire. Finally, Gottman and Levenson (1999b) found that cognitive variables coded from the Oral History Interview (Gottman, 1999) significantly predicted the quality of marital interactions 1 year later.

The clinical literature reflects considerable divergence on the issue of whether initial assessment of couple distress should be conducted with partners conjointly or should also include individual interviews with partners separately. Arguments for the latter include considerations of both veridicality and safety, particularly when assessing such sensitive issues as partner violence or substance abuse. Research indicates that couples experiencing domestic violence often do not disclose a partner's violent behavior in early interviews because of embarrassment, minimization, or fear of retribution (Ehrensaft & Vivian, 1996). Moreover, risks of retaliatory aggression against one partner by disclosing the other's violence in conjoint interview argue for the importance of conducting inquiries concerning partner violence in individual interviews. A common protocol combining conjoint and individual interviews for assessing couple distress involves an initial meeting with the couple together, followed by separate sessions with each partner individually, and then an additional conjoint meeting to provide a preliminary formulation and outline treatment alternatives (Karpel, 1994); within this format, the lengths of conjoint and individual components of interviews may vary so that the entire process is completed in one session (typically ranging from 1 to 2 hr) or across several sessions. Arguments against individual interviews when assessing couple distress emphasize potential difficulties in conjoint therapy if one partner has disclosed information to the therapist about which the other partner remains uninformed. Hence, if separate interviews are conducted with partners as a prelude to conjoint couple therapy, the interviewing clinician needs to be explicit with both partners ahead of time regarding conditions under which information disclosed by one partner will be shared with the other and any criteria for selecting among individual, conjoint, or alternative treatment modalities. We refer readers to extended discussions of the complex issues involved in assessing and treating partner violence available elsewhere (e.g., Aldarondo & Straus, 1994; Bograd & Mederos, 1999; O'Leary & Maiuro, 2001; Rathus & Feindler, 2004).

In sum, the research literature suggests that a couple interview can provide valid and clinically useful measures of many constructs relevant to couple distress. However, the dearth of psychometric studies on structured and comprehensive couple interviews makes it impossible to draw inferences about their sources of error variance and the best query and response formats for obtaining optimally valid and useful measures. Furthermore, there are no data on many of the important functions of the interview, such as strengthening client adherence to treatment protocols.

Observational Methods

An empirically driven couple assessment would seem empty without an observation of the couple's interaction. Assessment is both a hypothesis-generating and hypothesis-testing enterprise, and ABO is a powerful tool on both accounts. Because partners frequently present for treatment together, clinicians have the rare opportunity to assess the reciprocal social determinants of problem behaviors without venturing outside the therapy office.

Like interviews and self-report methods, ABO describes a method of data collection; it is not a measure in its own right. Using ABO in empirically driven assessment requires five key knowledge domains, all of which are subsumed under the strategy labeled by Haynes and O'Brien (2000, p. 89) as a "scholarly, empirical, hypothesis-testing approach to assessment." First, the assessor should know why ABO, generally, may be a useful assessment method. Second, the assessor should know the classes of behavior for which ABO measurement tools exist. Third, the assessor should know the results of research literature on the reliability and validity of ABO measurement tools. Further, the assessor should understand both the contexts in which the tools have been used and the conditional nature of validity. Fourth, the assessor should know how to use ABO in a clinical assessment. Finally, the assessor should know the limitations of ABO generally and ABO in standard clinical practice specifically.

Why ABO Is a Useful Couple Assessment Method

Heyman and Slep (2004, p. 162) defined ABO as involving ". . . a situation designed by, manipulated by, or constrained by an [assessor] that elicits a measured behavior of interest Observed behaviors [can] comprise both verbal and nonverbal emissions (e.g., motor actions, verbalized attributions, observable facial reactions)." Three elements of the definition should be highlighted: First, the analog (vs. purely naturalistic) nature of the observation involves the assessor structuring the situation, allowing the assessor to "stack the deck" to make it more likely that the behaviors (or functional relations) of interest will occur when the assessor can see them (Haynes, 2001). Naturally, the analog situation (or situations; e.g., discussing a conflict, discussing personal vulnerabilities) will be selected on the basis of hypotheses that the assessor wishes to test. Second, behaviors and behavioral sequences to be measured are already operationalized; the assessor is applying, not creating, the measurement tool. Finally, the range of behaviors for which tested observation tools exist is quite broad and is increasing at a healthy clip (see Kerig & Baucom, 2004).

ABO can be a useful and valid assessment tool because, depending on how it is applied, it minimizes inferences needed to assess behavior, it can facilitate formal or informal functional analysis, it can provide the assessor with experimental control of situational factors (thus helping to isolate the determinants of behavior), it can provide an additional method of assessment in a multimethod strategy (e.g., questionnaires, interviews, observation), and it can facilitate the observation of otherwise difficult to observe behaviors (Haynes & O'Brien, 2000; Heyman & Slep, 2004).

Classes of Couples' Behavior for Which ABO Measurement Tools Exist

A recent edited volume (Kerig & Baucom, 2004) provides detailed chapters on 15 coding systems; Heyman's (2001) psychometric-oriented review of couple coding systems included over 25 more. Space limitations preclude a detailed summary of these two works. However, in our estimation, there are six major a priori classes of behaviors measured in these systems: affect (e.g., Specific Affect Coding System, Shapiro & Gottman, 2004), communication behavior/affect (e.g., Kategoriensystem für Partnerschaftliche Interaktion, Hahlweg, 2004; Hahlweg et al., 1984; Rapid Marital Interaction Coding System, Heyman, 2004), everyday behaviors (e.g., Turning Toward vs. Turning Away; Driver & Gottman, 2004), power (e.g., System for Coding Interactions in Dyads, Malik & Lindahl, 2004), problem solving (e.g., Communication Skills Test; Floyd, 2004), and support/intimacy (e.g., Social Support Interaction Coding System, Pasch, Harris, Sullivan, & Bradbury, 2004; Social Support Behavior Code, Suhr, Cutrona, Krebs, & Jensen, 2004).

Psychometrics of Couple ABO Measurement Tools

Interobserver agreement and reliability. The interested reader is referred to Heyman (2001) and Kerig and Baucom (2004) for comprehensive information on psychometrics of couple ABO measurement tools. In short, nearly all published systems have adequate interobserver agreement when all codes are considered at once. Reporting agreement in this manner makes it difficult to discern whether the requisite agreement at the level of analysis was satisfactory (Heyman, 2001). Fortunately, most studies published or conducted within the last 5–10 years have included interobserver agreement metrics at the construct level, and the level of agreement is adequate (see Kerig & Baucom, 2004). Reliability (temporal stability) has never been fully tested, although a generalizability study by Wieder and Weiss (1980) indicates considerable variability across interactions, and studies by Christensen and Heavey (1990) and Heavey, Layne, and Christensen (1993) indicate that behavioral stability is influenced by which partner is pursuing change in the conversation. The two studies examining stability of observations across 4–5 years (Gottman & Levenson, 1999a; Lord, 1999) reported conflicting results.

Validity. Before discussing validity, it is useful to remember that, as Haynes et al. (1995, pp. 239–241) noted, "[V]alidity is a state, not a trait, of an obtained assessment instrument score Statements such as ' . . . has been shown to be a reliable and valid assessment instrument' do not reflect the conditional nature of validity and are usually unwarranted." The strongest, most replicated discriminative validity findings in couples observation research can be summarized as follows:

Distressed partners, compared with nondistressed partners (a) are more hostile, (b) start their conversations more hostilely and maintain it during the course of the conversation, (c) are more likely to reciprocate and escalate their partners' hostility, (d) are less likely to edit their behavior during conflict, resulting in longer negative reciprocity loops, (e) emit less positive behavior, (f) suffer more ill health effects from their conflicts, and (g) are more likely to show demand → withdraw patterns. Furthermore, both partners in distressed relationships characterized by husband-to-wife aggression, compared

Table 2
Coding Constructs and Systems

Construct	Exemplar system	Sample code	Primary type of situation used	Level of coding
Affect	SPAFF (Shapiro & Gottman, 2004)	Criticism, contempt, defensiveness, stonewalling, affection, humor	Conflict, events of the day	Midlevel microanalytic
Communication behavior/affect	KPI (Hahlweg, 2004); RMICS (Heyman, 2004)	Hostility, devaluation of partner, withdrawal, acceptance, self-disclosure	Conflict, problem solving, hypothetical couple situations	Midlevel microanalytic
Everyday behaviors	Turning Toward Versus Turning Away (Driver & Gottman, 2004)	Bids, responses, shared moments, turning away, and rebids	Unstructured observed time together	Microanalytic
Power	System for Coding Interactions in Dyads (Malik & Lindahl, 2004)	Coerciveness, attempts to control, negative escalation, balance of power	Conflict	Global
Problem solving	Communication Skills Test (Floyd, 2004)	Five-point scale (for each talk turn) ranging from very negative to very positive	Conflict	Midlevel microanalytic rating scale
Support/intimacy	Support Behavior Code (Suhr, Cutrona, Krebs, & Jensen, 2004), SSICS (Pasch, Harris, Sullivan, & Bradbury, 2004)	Understanding, empathy, reassurance, validation, positive instrumental behavior	Discloser/listener roles assigned for nonconflictual personal problem discussion	Microanalytic

Note. SPAFF = Specific Affect Coding System; KPI = Kategoriensystem für Partnerschaftliche Interaktion; RMICS = Rapid Marital Interaction Coding System; SSICS = Social Support Interaction Coding System.

with distressed/nonaggressive relationships, are more hostile and reciprocate hostility more. (Heyman, 2001, p. 6)

Specific measures have also demonstrated convergent validity (with questionnaires and with other systems) and sensitivity to change following couple therapy (see validity sections of chapters in Kerig & Baucom, 2004). However, given the conditional nature of validity, one must be cautious about inferring that research-based coding systems are valid measures when used in standard clinical assessment. We return to this caveat below.

Generalizability across populations. Grossly similar findings have been reported in samples from Australia, Canada, Germany, Holland, Spain, and the United States (Heyman, 2001) and in both heterosexual and same-sex couples (Kline et al. 2004; Shapiro & Gottman, 2004). Although these systems obtained good interobserver agreement for couples from diverse racial/ethnic and economic backgrounds, generalizability across such backgrounds has not been systematically tested.

How to Use ABO in a Clinical Assessment of Couples

Case formulation, or conceptualization, is defined as “a general model . . . to understand problems and generate solutions to them, based on this understanding, in a coherent, systematic way” (Persons, 1989, p. xiii). In other words, the key purpose of pretreatment assessment is formulating and testing initial hypotheses about the key causal and maintaining factors of the presenting problem (or problems; Haynes & O’Brien, 2000). Collecting communication samples is an important part of couples’ clinical assessment because “[c]ommunication is the common pathway to relationship dysfunction because it is the common pathway for getting what you want in relationships. Nearly all relationship-relevant conflicts, emotions, and neuroses are played out via observable communication—either verbally or nonverbally” (Heyman, 2001, p.

6). However, if questionnaire or interview assessments suggest that an interactive task may place one or both partners in danger (e.g., if there is a history of serious physical or emotional abuse, indications of severe power or control dynamics, or threats conveyed to the assessor), ABO would be contraindicated.

If it seems reasonable that it is safe to proceed, then the clinician should hypothesize which classes of behaviors seem most highly connected to the target problems. Furthermore, unless the assessor can rule out a plausible connection between conflict communication and the couple’s problems, we recommend that a conflict communication ABO be collected.

The clinician should then become familiar with the coding systems that assess that class of behaviors and with procedures for setting up the analog situations that best elicit behaviors of that class (cf., Heyman, 2001; Heyman & Slep, 2004; Kerig & Baucom, 2004). We also recommend, wherever possible, that the ABO be videorecorded so that the sample can be reviewed later with an eye toward a class of behaviors other than what was the assessor’s primary focus during the in vivo ABO. Classes of specific behaviors for which ABO procedures have been developed, exemplar coding systems, sample codes and primary types of situations, and level of coding used within that system are listed in Table 2.² Figure 1 displays a flowchart for employing ABO in the clinical assessment of couples. Upon reviewing the couple’s history of

² Because there are more than 40 coding systems summarized in the Heyman (2001) article and Kerig and Baucom (2004) book, we will limit our exemplars to (a) the most widely used systems in each class that (b) have demonstrated interrater agreement and indicators of validity and (c) comprise a manageable number of codes for assessors. Thus, widely used ultramicroanalytic systems, such as the Marital Interaction Coding System (Heyman, Weiss, & Eddy, 1995), that have dozens of codes are too unwieldy for everyday use and are not used as exemplars.

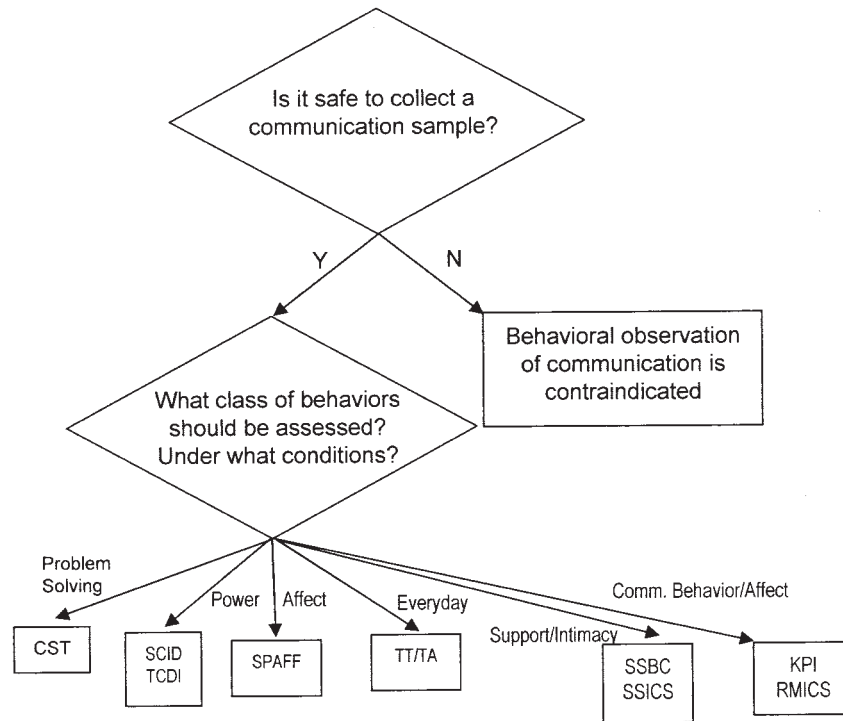


Figure 1. Flowchart for using observation in empirically driven clinical assessment. Comm. = communication; CST = Communication Skills Test; SCID = System for Coding Interactions in Dyads; TCDI = Thematic Coding of Dyadic Interaction (Vivian, Langhinrichsen-Rohling, & Heyman, 2004); SPAFF = Specific Affect Coding System; TT/TA = turning toward/turning away; SSBC = Social Support Behavior Code; SSICS = Social Support Interaction Coding System; KPI = Kategoriensystem für Partnerschaftliche Interaktion; RMICS = Rapid Marital Interaction Coding System.

partner aggression, if the clinician deems it safe to collect a behavioral sample of the couple's communication, then decisions should be reached regarding which classes of behaviors to assess (e.g., problem solving, social support) and which observational systems best lend themselves to coding these behaviors.

Limitations of ABO

One limitation of ABO research with couples involves such findings being culture bound. Malik and Lindahl (2004) reported that different variables discriminated problematic and nonproblematic European American and Latino families. Thus, validity results reported for many systems (and thus the knowledge that they seem to impart about relationships) may not generalize beyond European or European-descended middle class volunteers and may weaken across time as cultural expectations for relationship behavior shift.

Moreover, concerns have been raised about ABO's somewhat circumscribed clinical utility (e.g., Mash & Foster, 2001). All coding systems require scores of hours of observer training to reach adequate levels of interobserver agreement. Even after observers are certified as reliable, a great deal of energy is expended to maintain reliability (e.g., weekly meetings with regular feedback on agreement). This "coder drift" follows a natural law (i.e., entropy) and thus is more reflective of the difficulty in maintaining a singular viewpoint than it is of some kind of intellectual failure

on the part of coding-system designers or observers. Thus, even if clinicians expended a great deal of time learning a system to the point of mastery (i.e., meeting the reliability criterion), their reliability would naturally decay without ongoing efforts to maintain agreement. Such a requirement is likely not reasonable for most clinicians. Alternatively, a specialized assessment center could analyze observations (e.g., Mash & Hunsley, 2004). Putting aside the daunting logistic and financial considerations, such ABO assessment would have to prove its worth by significantly improving clinicians' functional analyses and thus improve treatment gains (Strosahl & Robinson, 2004).

Thus, routine use of empirically based couple assessment using ABO in clinical settings remains elusive. Psychometric hurdles (i.e., interobserver agreement, reliability, validity) probably will continue to make even clinician-administered, empirically driven couple ABO unachievable (Mash & Hunsley, 2004). However, empirically informed use of ABO should be standard in assessing couple distress in clinical as well as research settings. As noted above, empirically informed observers can use ABO as part of a multimethod assessment strategy in an effort to test clinical hypotheses informally via (a) familiarity with the findings about healthy and dysfunctional relationships, (b) formation of hypotheses as part of initial assessment activities, and (c) choosing of ABO methods (i.e., tasks and measures) that are appropriate to test the hypotheses. On the basis of findings from ABO research with

couples, Heyman (2001) suggested that clinicians use ABO in assessing couple distress to address the following:

1. How does the conversation start? Does the level of anger escalate? What happens when it does? Does the couple enter repetitive negative loops?
2. Do partners indicate afterward that what occurred during the conversations is typical? Is their behavior stable across two or more discussions?
3. Do partners' behaviors differ when it is her topic versus his? Do they label the other person or the communication process as the problem?

Self- and Other-Report Methods

The rationale underlying self-report methods³ in couple assessment is that such methods (a) are convenient and relatively easy to administer; (b) are capable of generating a wealth of information across a broad range of domains and levels of functioning germane to clinical assessment or research objectives including those listed in Table 1; (c) lend themselves to collection of data from large normative samples that can serve as a reference for interpreting data from individual respondents; (d) allow disclosure about events and subjective experiences respondents may be reluctant to discuss with an interviewer or in the presence of their partner; and (e) can provide important data concerning internal phenomena opaque to observational approaches including thoughts and feelings, values and attitudes, expectations and attributions, and satisfaction and commitment.

However, the limitations of traditional self-report measures also bear noting. Specifically, data from self-report instruments can (a) reflect bias in self- and other presentation in either a favorable or unfavorable direction, (b) be affected by differences in stimulus interpretation and errors in recollection of objective events, (c) inadvertently influence respondents' nontest behavior in unintended ways, and (d) typically provide few fine-grained details concerning moment-to-moment interactions compared with ABO. Because of their potential advantages, and despite their limitations, self-report techniques of couple and family functioning have proliferated, with published measures numbering well over 1,000 (Touliatos, Perlmutter, Straus, & Holden, 2001). However, relatively few of these measures have achieved widespread adoption. Chun, Cobb, and French (1975) found that 63% of measures they reviewed had been used only once, with only 3% being used 10 times or more. Fewer than 40% of marital and family therapists regularly use any standardized instruments (Boughner, Hayes, Bubenzer, & West, 1994). Contributing to these findings is the inescapable conclusion that the majority of measures in this domain demonstrate little evidence regarding the most rudimentary psychometric features of reliability or validity, let alone clear evidence supporting their clinical utility (Snyder & Rice, 1996).

We describe below a small subset of self-report instruments selected on the basis of their representativeness across behavioral, cognitive, and affective domains of couples' interactions, their potential clinical utility, and at least moderate evidence of their reliability and validity. In some domains (e.g., relationship cognitions) existing measures are few, and we have drawn on recom-

mendations outlined in previous reviews (cf., Epstein & Baucom, 2002; Sayers & Sarwer, 1998). We include some measures because of unique and contrasting functions they serve in clinical or research settings—for example, broad multidimensional measures for differentiating sources and levels of relationship distress useful for treatment planning and outcomes assessment versus brief global screening measures of couple distress useful in research or primary care settings. In a very few instances, we include recent measures that comprise adaptations or extensions of previous instruments, but for which few psychometric data are available; in each case we note this. Although we strive to identify representative measures across domains yet constrain the number of measures presented, our selections have admittedly been influenced by our own experiences in both clinical and research applications.

More comprehensive bibliographies of self-report couple and family measures are available elsewhere (e.g., Corcoran & Fischer, 2000; Davis, Yarber, Bauserman, Schreer, & Davis, 1998; Fredman & Sherman, 1987; Grotevant & Carlson, 1989; Jacob & Tennenbaum, 1988; L'Abate & Bagarozzi, 1993; Touliatos et al., 2001). We also refer readers to recent texts on couple therapy that emphasize the clinical use of self-report measures as an integral component of planning and evaluating couple interventions (cf., Epstein & Baucom, 2002; Gottman, 1999). Although some of the measures described here have been published commercially, and others are included in their entirety in journal or book publications, many are available only by requesting them from their original author. The Touliatos et al. (2001) handbook comprises the best resource regarding the content, format, and availability of couple and family measures.

Measures of Behavior

Although distinctions among measures of behavior, cognition, and affect are imperfect, we focus here on measures purporting to assess couples' behavior exchanges, including communication, verbal and physical aggression, and sexual intimacy. One of the earliest and most widely used measures of couples' behavior is the SOC (Birchler et al., 1975), a list of 400 discrete behaviors divided on an a priori basis into 12 categories such as affection and physical intimacy, companionship, communication, parenting, finances, and division of household responsibilities. Although specific administration instructions may vary, each individual is asked to complete the checklist covering a specific time period (e.g., the previous 24 hr), indicating which behaviors their partner had emitted and whether these were experienced as pleasing or displeasing. As a clinical tool, the SOC generates menus of individual reinforcers and has the potential to delineate relative strengths and weaknesses in the relationship, transforming diffuse negative complaints into specific requests for positive change. A brief adaptation of the SOC asks respondents to provide summary satisfaction ratings for each of the 12 SOC categories (O'Leary, 1987).

Other behavioral measures attempt to identify specific areas of desired change, the amount and direction of change desired, the

³ In this article we use *self-report* to refer specifically to written quantitative and qualitative responses to questionnaire measures, as distinguished from verbal reports about oneself or another person obtained in interview.

congruence of desired change across partners, and individuals' accuracy in perceiving their partner's wishes. The older of these two, the ACQ (Weiss & Birchler, 1975), presents each partner with two identical lists of 34 specific behaviors (e.g., helping with housework or spending more time in outside activities) and asks each individual to indicate whether they would like their partner to increase or decrease that behavior and whether an increase or decrease in his or her own rate of that behavior would be pleasing to the partner. Scoring algorithms for the ACQ have been described for evaluating overall levels of desired change, congruence of partners' desired change in specific behaviors, as well as perceptual accuracy of each individual's understanding of their partner's wishes. A recent alternative to the ACQ, Gottman's (1999) Areas of Change Checklist, adopts a simpler approach in listing 36 potential relationship problems and asking respondents to rate the level of desired change for each item on a 5-point scale. Although simpler to complete than the ACQ, no data regarding psychometric characteristics of the measure have been reported.

A variety of self-report measures of communication have been developed, several of which are described in an excellent review by Sayers and Sarwer (1998). The CPQ (Christensen, 1987) was designed to measure the temporal sequence of couples' interactions by soliciting partners' perceptions of their communication patterns before, during, and following conflict. Scores on the CPQ can be used to assess characteristics of the demand → withdraw pattern frequently observed among distressed couples. An alternative measure of couples' communication is the SCI (Metz, 1993), a 126-item inventory that elicits individuals' descriptions of their own behavior in response to a conflict situation as well as thoughts and perceptions of their partner's behavior. Scores on the SCI permit comparisons of partners' appraisals with each other as well as comparisons with a standardization sample along dimensions reflecting frequency, intensity, and attributions regarding responsibility for relationship conflicts.

Assessing relationship aggression by self-report measures assumes particular importance because of some individuals' reluctance to disclose the nature or extent of such aggression during an initial conjoint interview. By far, the most widely used measure of couples' aggression is the CTS. The original CTS (Straus, 1979) included 19 items assessing three modes of conflict resolution: reasoning, verbal aggression, and physical aggression. The revised instrument (CTS2; Straus, Hamby, Boney-McCoy, & Sugarman, 1996) adds scales of sexual coercion and physical injury as well as additional items to better differentiate between minor and severe levels of verbal and physical aggression. Complementing the CTS is the Psychological Maltreatment of Women Inventory (Tolman, 1989) and its gender-neutral version (Kasian & Painter, 1992) containing 58 items designed to assess such components of emotional abuse as demands for subservience, isolation from external support, withholding of emotional resources, verbal attacks, and related degrading behaviors. An additional measure of relationship aggression, the Aggression subscale of the MSI-R (Snyder, 1997), comprises 10 items reflecting psychological and physical aggression experienced from one's partner. Advantages of the Aggression subscale as a screening measure include its relative brevity and its inclusion in a multidimensional measure of couples' relationships (the MSI-R) described below. More extensive discussions of clinical strategies and specific measures for assessing

physical and psychological aggression in relationships are noted earlier in this article.

Similar to individuals with issues of aggression, some individuals may be reluctant to disclose intimate details of their sexual relationship during an initial interview. Numerous measures of sexual attitudes, behaviors, and conflicts have been developed (Davis et al., 1998). Two more widely used self-report techniques in this domain include the Sexual Interaction Inventory (SII; LoPiccolo & Steger, 1974) and the Derogatis Sexual Functioning Inventory (DSFI; Derogatis, Lopez, & Zinzeletta, 1988; Derogatis & Melisaratos, 1979). The SII is a 102-item measure that asks the individual to rate the frequency of activity and levels of satisfaction, both real and ideal for both self and partner, across 17 behaviors ranging from intercourse to nudity and nonsexual physical intimacy. By comparison, the DSFI includes 254 items composing 10 scales reflecting such areas as sexual knowledge, range of sexual experiences, sexual attitudes and drive, as well as psychological symptoms in nonsexual domains. Although concerns have been raised in the literature regarding the veridicality of self-reports, particularly those in regard to specific sexual practices (cf., McConaghy, 1998), both the SII and DSFI have garnered support for their reliability and discriminative validity, and responses to either measure can be used to introduce sensitive issues in clinical interview.

Measures of Cognition

Earlier we noted the importance of evaluating couples' assumptions, standards, attentional sets, expectancies, and attributions for relationship events. Several self-report measures have been developed to assist in this process. The DAI (Baucom et al., 1989) is a 24-item measure that asks respondents to imagine hypothetical relationship events and then, for each event, generate attributions for their partner's behavior in that situation with regard to (a) source of influence (self, partner, or external factors), (b) stability or instability of causal factors, and (c) their specificity or globality. The DAI is designed to assist clinicians in identifying and modifying dysfunctional attributional sets contributing to subjective negativity surrounding specific relationship events.

An alternative measure, the Relationship Attribution Measure (Fincham & Bradbury, 1992), also presents hypothetical situations but asks respondents to generate responsibility attributions indicating the extent to which the partner intentionally behaved negatively, was selfishly motivated, and was blameworthy for the event. A third attributional measure, the Marital Attitude Survey (Pretzer et al., 1992), elicits attributions along six dimensions reflecting causal influence from one's own behavior or personality, the partner's behavior or personality, and attributions regarding the partner's malicious intent or lack of love. The moderating role of these dimensions was demonstrated in a study in which the relation of marital distress to depressive symptomatology was greater for wives who attributed marital difficulties to their own behavior than for wives who attributed marital difficulties to their husbands' behavior (Heim & Snyder, 1991).

Separate from attributional measures have been those examining unrealistic relationship assumptions or standards. An early measure in this domain, the Relationship Beliefs Inventory (Eidelson & Epstein, 1982), assesses five dysfunctional ideas about marriage—for example, that disagreements are necessarily destructive

or that partners should know each other's feelings and thoughts without asking. A more recent 60-item measure, the ISRS (Baum et al., 1996), assesses three dimensions reflecting (a) partners' beliefs regarding optimal levels of independence versus sharing of time, interests, and activities; (b) standards regarding levels of power or control that should be exercised by each partner; and (c) expectations for how much each partner should contribute to the relationship in terms of emotional and behavioral investment. Scores on the ISRS can be used to guide clinical interventions; for example, couples reporting fewer individual boundaries, egalitarian decision making, and high relationship investment were more likely to have higher relationship satisfaction.

Measures of Relationship Sentiment

Measures of relationship satisfaction and global affect abound. The two oldest and most widely used are the Locke–Wallace Marital Adjustment Test (MAT; Locke & Wallace, 1959) and the DAS (Spanier, 1976). The MAT is a 15-item questionnaire that asks partners to rate their overall happiness in their relationship as well as their extent of agreement in key areas of interaction. Displacing the MAT as the most frequent global measure of relationship satisfaction is the DAS, a 32-item instrument purporting to differentiate among four related subscales reflecting cohesion, satisfaction, consensus, and affectional expression. For abbreviated screening measures of couple distress, several alternatives are available. The Relationship Satisfaction Scale (Burns & Sayers, 1992) is a 13-item Likert-type measure that assesses satisfaction in such areas as handling of finances and degree of affection and caring. The QMI (Norton, 1983) is a 6-item Likert-type measure asking respondents to rate their overall level of marital happiness and the accuracy of additional descriptors of overall relationship stability and accord. An even shorter measure, the KMSS (Schumm et al., 1986) includes three Likert items assessing satisfaction with marriage as an institution, the marital relationship, and the character of one's spouse. A brief (7-item) version of the DAS has also been developed and found to be psychometrically sound (Hunsley, Best, Lefebvre, & Vito, 2001). In general, abbreviated scales of global relationship satisfaction are adequate as initial screening measures in primary care or general psychiatric settings but, because of their brevity, lack the ability to distinguish reliably among finer gradations of relationship distress among partners presenting for couple therapy. Moreover, the relatively homogeneous content of brief measures of relationship satisfaction, although potentially a strength in terms of interpretive clarity, potentially limits the ability to detect couple distress provided by somewhat more heterogeneous but longer measures.

There is considerable convergence across measures of relationship affect purporting to assess such constructs as marital quality, satisfaction, adjustment, happiness, cohesion, consensus, intimacy, and the like, with correlations between measures often approaching the upper bounds of their reliability. Differentiation among such constructs at a theoretical level often fails to achieve the same operational distinction at the item-content level (cf., Fincham & Bradbury, 1987, for an excellent discussion of this issue). Moreover, factor analytic studies of multiscale measures often fail to support purported factor structure and differentiations at the subscale level; for example, factor analyses have failed to replicate the original four subscales of the DAS (Crane, Busby, & Larson,

1991), although such analyses have generally affirmed the multidimensionality of the DAS. Hence, selection among such measures should be guided by careful examination of item content and empirical findings regarding both convergent and discriminant validity as well as by such practical considerations as use of a screening device for identifying couple distress versus planning specific treatment components.

Finally, although not specifically comprising measures of relationship sentiment, recent measures grounded in attachment theory will likely find increased use both in clinical and research contexts as the relevance of attachment for adult relationships gains in popularity and empirical investigations. The linkages of self-report measures of adult attachment to observational and other measures of individual and relationship functioning have recently been reviewed by Shaver and Mikulincer (2004).

Multidimensional Measures

Well-constructed multidimensional measures of couples' interactions have the potential to discriminate among various sources of relationship strength, conflict, satisfaction, and goals. Although some of the measures described earlier provide scores on two or more scales, the facets of relationship functioning targeted by those respective measures tend to be restricted to one or two specific domains of couple distress. For example, although the SII facilitates differentiating potential sources of couple distress across specific components of physical intimacy, its length (102 items) likely restricts its clinical usefulness to couples for whom sexual difficulties have already been identified as a significant concern. By contrast, several measures have been developed for assessing couple distress across multiple dimensions of partners' interaction, so that identified problem areas may then be targeted for more detailed assessment with clinical interview, ABO, or selected measures emphasizing that specific domain. Among such measures obtaining fairly widespread use are the PREPARE and ENRICH inventories by Fowers and Olson (1989, 1992), developed for use with premarital and married couples, respectively. Both of these measures include 125 items rated on a 5-point scale and assessing relationship accord in such domains as communication, conflict resolution, the sexual relationship, and finances. A computerized interpretive report identifies areas of "strength" and "potential growth" and directs respondents to specific items reflecting potential concerns.

Also widely used in both clinical and research settings is the MSI-R, a 150-item inventory designed to identify both the nature and intensity of relationship distress in distinct areas of interaction. The MSI-R includes 2 validity scales, 1 global scale, and 10 specific scales assessing relationship satisfaction in such areas as affective and problem-solving communication, aggression, leisure time together, finances, the sexual relationship, role orientation, family of origin, and interactions regarding children. More than 20 years of research have supported the reliability and construct validity of the MSI-R scales (cf., Snyder & Aikman, 1999). Recent studies suggest the potential utility of Spanish and German adaptations of the MSI-R for cross-cultural application with both clinic and community couples (Snyder et al., 2004) as well as use of the original English version with nontraditional (e.g., gay and lesbian) couples (Means-Christensen, Snyder, & Negy, 2003).

Conclusions and Recommendations

Assessment of couple distress shares basic principles of assessing individuals—namely, that (a) the content of assessment methods be empirically linked to target problems and constructs hypothesized to be functionally related, (b) selected assessment methods demonstrate evidence of reliability and validity, and (c) findings be linked within an overarching theoretical or conceptual framework to the presumed causes of difficulties as well as to clinical intervention or prevention. However, couple assessment differs from individual assessment in that couple assessment strategies (a) focus specifically on relationship processes and the interactions between individuals, (b) provide an opportunity for direct observation of target complaints involving communication and other interpersonal exchange, and (c) must be sensitive to potential challenges unique to establishing a collaborative alliance when highly distressed or antagonistic partners are assessed, particularly in a conjoint context. Similar to the assessment process itself, our review of strategies for assessing couple distress has been necessarily selective, emphasizing dimensions empirically related to couple distress, identifying alternative methods for obtaining relevant assessment data, and highlighting specific techniques within each method. From this review, several recommendations regarding clinical assessment and further research can be extracted.

Recommendations for Assessing Couple Distress

Assessment strategies and specific methods for assessing couple distress will necessarily be tailored to partners' unique constellation of presenting difficulties as well as to specific resources of both the couple and the assessor. However, regardless of the specific context, the following recommendations for assessing couple distress generally apply.

1. Given empirical findings linking couple distress to individual disorders and their respective impact in moderating treatment outcome, assessment of couple functioning should be standard practice when treating individuals. Screening for couple distress when assessing individuals may involve a brief interview format shown to relate to relevant indicators of couple interactions (e.g., Forthofer et al., 1996) or a brief self-report measure exhibiting prior evidence of discriminative validity (e.g., the QMI, KMSS, or short-form DAS). Similarly, when treating couples, partners should be screened for individual emotional or behavioral difficulties potentially contributing to, exacerbating, or resulting in part from couple distress.

2. Assessment should logically progress from identifying relationship concerns at the broader construct level to examining more specific facets of couple distress and its correlates with a finer-grained analysis. The specific assessment methods described in this review vary considerably in their overall breadth or focus within any specific construct domain and, hence, will vary both in their applicability across couples and their placement in a sequential exploratory assessment process.

3. Within clinical settings, certain domains should always be assessed with every couple either because of their robust linkage to relationship difficulties (e.g., communication processes involving emotional expressiveness and decision making) or because the specific behaviors, if present, have particularly adverse impact on couple functioning (e.g., physical aggression or substance abuse).

4. Couple assessment should integrate findings across multiple assessment methods. Multiple indicators of relationship functioning should be used to reach clinical judgments in any domain, because any single measure or indicator can reflect multiple sources of error. Behavior, affect, and cognition should all be assessed. Self-reports and informal observational data acquired from interview should be complemented by structured ABO tailored to relationship tasks and classes of behavior hypothesized to be most highly linked to target problems.

5. Self- and other-report measures may complement findings from interview or behavioral observation in generating data across diverse domains both central or conceptually related to the couple's difficulties or across those domains potentially more challenging to assess because of their sensitive nature or their not being amenable to direct observation. However, special caution should be exercised when adopting self- or other-report measures in assessing couple distress. Despite their proliferation, most measures of couple functioning described in the literature have not undergone careful scrutiny of their underlying psychometric features. Among those instruments for which some evidence concerning reliability and validity has been garnered, evidence often exists only for overall scores and not at the level of subscales or smaller units of analysis at which interpretations may be made.

6. At the same time, assessment of couple distress should be parsimonious. This objective can be facilitated by choosing evaluation strategies and modalities that complement each other and by following a sequential approach that uses increasingly narrow-band measures to target problem areas that have been identified by other assessment techniques. Throughout our review, we have emphasized respective strengths and limitations of the clinical interview, behavioral observation, and self- and other-report methods. However, empirical findings regarding the incremental utility of specific measures within method or complementary measures across method—either in enhancing treatment effectiveness or in predicting discrete relationship events—are generally lacking.

7. Psychometric characteristics of any assessment technique—whether from interview, ABO, or self-report measure—are conditional upon the specific population and purpose for which that assessment method was developed. Given that nearly all measures of couple distress were developed and tested on White middle-class married couples, their relevance to, and utility for, assessing ethnic couples, gay and lesbian couples, and low-income couples is unknown. This caveat extends to content- as well as criterion-related validity. For example, Haynes and colleagues (1992) found that traditional measures of marital quality failed to assess important factors related to marital satisfaction for older persons (e.g., health), included nonrelevant content (e.g., child rearing), and confounded or misconstrued other factors (e.g., affection vs. sexual exchanges). Hence, any assessment measure demonstrating evidence of validity with some couples may not be valid, in part or in whole, for any given couple, further underscoring the importance of drawing upon multiple indicators across multiple methods for assessing any specific construct.

Recommendations for Further Research

Future directions for assessment research germane to the field generally also apply to research in assessing couple distress specifically, including the need for greater attention to (a) psychomet-

ric underpinnings of various measurement methods and instruments; (b) factors moderating reliability and validity across populations differing in sociocultural characteristics as well as in clinical functioning; (c) the assessment process including initial articulation of assessment goals, selection of assessment method and instruments, and methods of interpreting data and providing feedback; and (d) the functional utility of assessment findings in enhancing treatment effectiveness (Hayes, Nelson, & Jarrett, 1987).

In considering the implications of these directives for assessing couple distress, considerably more research is needed before a comprehensive empirically based couple assessment protocol can be advocated. For example, despite the ubiquitous use of couple assessment interviews, virtually no research has been conducted to assess their psychometric features. Observational methods, although a rich resource for generating and testing clinical hypotheses, are less frequently used and present significant challenges to their reliable and valid application in everyday practice. Questionnaires, despite their ease of administration and potential utility in generating a wealth of data, frequently suffer from inadequate empirical development and, at best, compose only part of a multimethod assessment strategy.

We recommend, as a research roadmap, that clinical researchers consider adapting the National Institutes of Health stages of intervention research cycle (Mrazek & Haggerty, 1994). Stage I involves identifying the disorder and measuring its prevalence. Despite being so basic a need, there currently exists no gold standard for discriminating distressed from nondistressed couples; the questionnaires typically used for such classifications are of limited sensitivity and specificity (Heyman et al., 2001). Stage II involves delineating specific risk and protective factors. As noted above, some replicated factors have been identified, although this research could be sharpened by defining groups more carefully (through Stage I above). Stage III (efficacy trials) involves tightly controlled trials of the efficacy of a multimethod assessment in clinical practice. Stage IV (effectiveness trials) would involve controlled trials of the outcome of this assessment in more real-world clinical environments. Only then would testing broad-scale dissemination (Stage V) of empirically based couple assessment be appropriate.

This research roadmap reflects an ambitious agenda unlikely to be met by any single investigator or group of investigators. However, progress toward evidence-based assessment of couple distress will be enhanced by research on specific components targeting more notable gaps in the empirical literature along the lines recommended below.

First, greater attention should be given to expanding the empirical support for promising assessment instruments already detailed in the literature than to the initial (and frequently truncated) development of new measures. Proposals for new measures should be accompanied by compelling evidence for their incremental utility and validity and a commitment to programmatic research examining their generalizability across diverse populations and assessment contexts.

Second, research needs to delineate optimal structured and semi-structured interview formats for assessing couples. Such research should address (a) issues of content validity across populations and settings, (b) organizational strategies for screening across diverse system levels and construct domains relevant to couple functioning (similar to branching strategies for the Structured Clinical Inter-

view for the *DSM* and related structured interviews for individual disorders), (c) relative strengths and limitations to assessing partners separately versus conjointly, (d) factors promoting the disclosure and accuracy of verbal reports, (e) relation of interview findings to complementary assessment methods (as in generating relevant tasks for ABO), and (f) the interview's special role in deriving functional-analytic case conceptualization.

Third, although laboratory-based behavioral observation of couple interaction has considerably advanced our understanding of couple distress, generalization of these techniques to more common clinical settings has lagged behind. Hence, researchers should develop more macrolevel coding systems for quantifying observational data that promote their routine adoption in clinical contexts while preserving their psychometric fidelity.

Fourth, research needs to attend to the influences of culture at several levels. First, there has been little attention to developing measures directly assessing domains specific to relationship functioning at the community or cultural level—for example, cultural standards or norms regarding emotional expressiveness, balance of decision-making influence, or boundaries governing the interaction of partners with extended family or others in the community. Hence, assessment of such constructs currently depends almost exclusively on the clinical interview, with no clear guidelines regarding either the content or format of questions. Second, considerably more research needs to examine the moderating effects of sociocultural factors on measures of couple functioning, including the impact of such factors as ethnicity, age, socioeconomic status, or sexual orientation. Third, work needs to proceed on adapting established measures to alternative languages. In the United States, the failure to adapt existing instruments to Spanish or to examine the psychometric characteristics of extant adaptations is particularly striking given that (a) Hispanics are the largest and fastest-growing ethnic minority group, and (b) among U.S. Hispanic adults ages 18–64, 28% have either limited or no ability to speak English (Snyder et al., 2004).

Adapting existing measures to alternative contexts (i.e., differing from the original development sample in language, culture, or specific aspects of the relationship such as sexual orientation) should proceed only when theoretical or clinical formulations suggest that the construct being measured does not differ substantially across the new application. Detailed discussions of both conceptual and methodological issues relevant to adapting tests to alternative languages or culture exist elsewhere (e.g., Butcher, 1996; Geisinger, 1994). Because clinicians and researchers may fail to recognize the inherent cultural biases of their conceptualization of couple processes, the appropriateness of using or adapting tests cross-culturally should be evaluated following careful empirical scrutiny examining each of the following:

1. Linguistic equivalence, including grammatical, lexical, and idiomatic considerations;
2. Psychological equivalence of items across the source and target cultures;
3. Functional equivalence, indicating the congruence of external correlates in concurrent and predictive criterion-related validation studies of the measure across applications; and

4. Scalar equivalence, ensuring not only that the slope of regression lines delineating test-criterion relations be parallel (indicating functional equivalence) but also that they have comparable metrics and origins (zero points) in both cultures.

Finally, research needs to examine the process, as well as the content, of couple assessment. For example, little is known regarding the impact of decisions about the timing or sequence of specific assessment methods, the role of the couple in determining assessment objectives, or the provision of clinical feedback on either the content of assessment findings or their subsequent effect on clinical interventions.

Although assessment of couples has shown dramatic gains in both its conceptual and empirical underpinnings over the past 25 years, much more remains to be discovered. Both clinicians and researchers need to avail themselves of recent advances in assessing couple distress and collaborate in promoting further development of empirically based assessment methods.

References

- Aldarondo, E., & Straus, M. (1994). Screening for physical violence in couple therapy: Methodological, practical, and ethical considerations. *Family Process, 33*, 425–439.
- Allgood, S. M., & Crane, D. R. (1991). Predicting marital therapy dropouts. *Journal of Marital and Family Therapy, 17*, 73–79.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Arellano, C. M., & Markman, H. J. (1995). The managing affect and differences scale (MADS): A self-report measure assessing conflict management in couples. *Journal of Family Psychology, 9*, 319–334.
- Baltas, Z., & Steptoe, A. (2000). Migration, culture conflict and psychological well being among Turkish-British married couples. *Ethnicity & Health, 5*, 173–180.
- Baucom, D. H., Epstein, N., & LaTaillade, J. J. (2002). Cognitive-behavioral couple therapy. In A. S. Gurman & N. S. Jacobson (Eds.), *Clinical handbook of couple therapy* (3rd ed., pp. 26–58). New York: Guilford Press.
- Baucom, D. H., Epstein, N., Rankin, L. A., & Burnett, C. K. (1996). Assessing relationship standards: The Inventory of Specific Relationship Standards. *Journal of Family Psychology, 10*, 72–88.
- Baucom, D. H., Sayers, S. L., & Duhe, A. (1989). Attributional styles and attributional patterns among married couples. *Journal of Personality and Social Psychology, 56*, 596–607.
- Baucom, D. H., Shoham, V., Mueser, K. T., Daiuto, A. D., & Stickle, T. R. (1998). Empirically supported couples and family therapies for adult problems. *Journal of Consulting and Clinical Psychology, 66*, 53–88.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory—II*. San Antonio, TX: Psychological Corporation.
- Bhugra, D., & De Silva, P. (2000). Couple therapy across cultures. *Sexual & Relationship Therapy, 15*, 183–192.
- Birchler, G. R., Weiss, R. L., & Vincent, J. P. (1975). Multimethod analysis of social reinforcement exchange between maritally distressed and nondistressed spouse and stranger dyads. *Journal of Personality and Social Psychology, 31*, 349–360.
- Bograd, M., & Mederos, F. (1999). Battering and couples therapy: Universal screening and selection of treatment modality. *Journal of Marital and Family Therapy, 25*, 291–312.
- Boughner, S. R., Hayes, S. F., Bubenzer, D. L., & West, J. D. (1994). Use of standardized assessment instruments by marital and family therapists: A survey. *Journal of Marital and Family Therapy, 20*, 69–75.
- Bradbury, T. N., & Fincham, F. D. (1990). Attributions in marriage: Review and critique. *Psychological Bulletin, 107*, 3–33.
- Bradbury, T. N., & Karney, B. R. (1993). Longitudinal study of marital interaction and dysfunction: Review and analysis. *Clinical Psychology Review, 13*, 15–27.
- Bradbury, T. N., Rogge, R., & Lawrence, E. (2001). Reconsidering the role of conflict in marriage. In A. Booth, A. C. Crouter, & M. Clements (Eds.), *Couples in conflict* (pp. 59–81). Mahwah, NJ: Erlbaum.
- Burns, D. D., & Sayers, S. L. (1992). *Development and validation of a brief relationship satisfaction scale*. Unpublished manuscript.
- Butcher, J. N. (1996). Translation and adaptation of the MMPI–2 for international use. In J. N. Butcher (Ed.), *International adaptations of the MMPI–2: A handbook of research and clinical applications* (pp. 26–43). Minneapolis: University of Minnesota Press.
- Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *Minnesota Multiphasic Personality Inventory—2: Manual for administration and scoring*. Minneapolis: University of Minnesota Press.
- Christensen, A. (1987). Detection of conflict patterns in couples. In K. Hahlweg & M. J. Goldstein (Eds.), *Understanding major mental disorder: The contribution of family interaction research* (pp. 250–265). New York: Family Process Press.
- Christensen, A., & Heavey, C. L. (1990). Gender and social structure in the demand/withdraw pattern of marital conflict. *Journal of Personality and Social Psychology, 59*, 73–81.
- Chun, K., Cobb, S., & French, J. R. P. (1975). *Measures for psychological assessment: A guide to 3,000 original sources and their applications*. Ann Arbor: University of Michigan, Survey Research Center of the Institute for Social Research.
- Cone, J. D. (1988). Psychometric considerations and the multiple models of behavioral assessment. In A. S. Bellack & M. Hersen (Eds.), *Behavioral assessment: A practical handbook* (3rd ed., pp. 42–66). New York: Pergamon Press.
- Cook, J., Tyson, R., White, J., Rushe, R., Gottman, J. M., & Murray, J. (1995). The mathematics of marital conflict: Qualitative dynamic mathematical modeling of marital interaction. *Journal of Family Psychology, 9*, 110–130.
- Corcoran, K., & Fischer, J. (2000). *Measures for clinical practice: A sourcebook: Vol. 1. Couples, families, and children*. New York: Free Press.
- Coyne, J. C. (1986). The significance of the interview in strategic marital therapy. *Journal of Strategic and Systemic Therapies, 5*, 63–70.
- Crane, D. R., Busby, D. M., & Larson, J. H. (1991). A factor analysis of the Dyadic Adjustment Scale with distressed and nondistressed couples. *American Journal of Family Therapy, 19*, 60–66.
- Cutrona, C. (1996). *Social support in couples: Marriage as a resource in times of stress*. Thousand Oaks, CA: Sage.
- Daiuto, A. D., Baucom, D. H., Epstein, N., & Dutton, S. S. (1998). The application of behavioral couples therapy to the assessment and treatment of agoraphobia: Implications of empirical research. *Clinical Psychology Review, 18*, 663–687.
- Davis, C. M., Yarber, W. L., Bauserman, R., Schreer, G., & Davis, S. L. (1998). *Handbook of sexuality-related measures*. Thousand Oaks, CA: Sage.
- Derogatis, L. R., Lopez, M. C., & Zinzeletta, E. M. (1988). Clinical applications of the DSFI in the assessment of sexual dysfunctions. In R. Brown & E. Roberts (Eds.), *Advances in the understanding and treatment of sexual disorders*. New York: PNA.
- Derogatis, L. R., & Melisaratos, N. (1979). The DSFI: A multidimensional measure of sexual functioning. *Journal of Sex and Marital Therapy, 5*, 244–281.
- Derogatis, L. R., & Savitz, K. L. (1999). The SCL-90-R, Brief Symptom Inventory, and matching clinical rating scales. In M. E. Maruish (Ed.),

- The use of psychological testing for treatment planning and outcomes assessment* (2nd ed., pp. 679–724). Mahwah, NJ: Erlbaum.
- Diener, E., Gohm, C. L., Suh, E., & Oishi, S. (2000). Similarity of the relations between marital status and subjective well-being across cultures. *Journal of Cross-Cultural Psychology, 31*, 419–436.
- Driver, J. L., & Gottman, J. M. (2004). Turning Toward Versus Turning Away: A coding system of daily interactions. In P. K. Kerig & D. H. Baucom (Eds.), *Couple observational coding systems* (pp. 209–226). Mahwah, NJ: Erlbaum.
- Durham, R. C., Allan, T., & Hackett, C. A. (1997). On predicting improvement and relapse in generalized anxiety disorder following psychotherapy. *British Journal of Clinical Psychology, 36*, 101–119.
- Ehrensaft, M., & Vivian, J. D. (1996). Spouses' reasons for not reporting existing physical aggression as a marital problem. *Journal of Family Psychology, 10*, 443–453.
- Eidelson, R. J., & Epstein, N. (1982). Cognitive and relationship maladjustment: Development of a measure of dysfunctional relationship beliefs. *Journal of Consulting and Clinical Psychology, 50*, 715–720.
- Epstein, N. B., & Baucom, D. H. (2002). *Enhanced cognitive-behavioral therapy for couples: A contextual approach*. Washington, DC: American Psychological Association.
- Fals-Stewart, W., Birchler, G. R., & O'Farrell, T. J. (2003). Alcohol and other substance abuse. In D. K. Snyder & M. A. Whisman (Eds.), *Treating difficult couples: Helping clients with coexisting mental and relationship disorders* (pp. 159–180). New York: Guilford Press.
- Fincham, F. D., & Bradbury, T. N. (1987). The assessment of marital quality: A reevaluation. *Journal of Marriage and the Family, 49*, 797–809.
- Fincham, F. D., & Bradbury, T. N. (1992). Assessing attributions in marriage: The Relationship Attribution Measure. *Journal of Personality and Social Psychology, 62*, 457–468.
- Finn, S. E., & Tonsager, M. E. (1997). Information-gathering and therapeutic models of assessment: Complementary paradigms. *Psychological Assessment, 9*, 374–385.
- First, M. B., Gibbon, M., Spitzer, R. L., & Williams, J. B. W. (1997). *Structured Clinical Interview for DSM-IV Axis I Disorders—Clinician version*. Washington, DC: American Psychiatric Association.
- Floyd, F. J. (2004). Communication Skills Test (CST) observational system for couples' problem-solving skills. In P. K. Kerig & D. H. Baucom (Eds.), *Couple observational coding systems* (pp. 143–158). Mahwah, NJ: Erlbaum.
- Forthofer, M. S., Markman, H. J., Cox, M., Stanley, S., & Kessler, R. C. (1996). Associations between marital distress and work loss in a national sample. *Journal of Marriage and the Family, 58*, 597–605.
- Fowers, B., & Olson, D. (1989). ENRICH marital inventory: A discriminant validity study. *Journal of Marital and Family Therapy, 15*, 65–79.
- Fowers, B., & Olson, D. (1992). Four types of premarital couples: An empirical typology based on PREPARE. *Journal of Family Psychology, 6*, 10–12.
- Fredman, N., & Sherman, R. (1987). *Handbook of measurements for marriage and family therapy*. New York: Brunner/Mazel.
- Geisinger, K. F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment, 6*, 304–312.
- Goering, P. N., Lancee, W. J., & Freeman, J. J. (1992). Marital support and recovery from depression. *British Journal of Psychiatry, 160*, 76–82.
- Gohm, C. L., Oishi, S., Darlington, J., & Diener, E. (1998). Culture, parental conflict, parental marital status, and the subjective well-being of young adults. *Journal of Marriage and the Family, 60*, 319–334.
- Gordis, E. B., Margolin, G., & John, R. S. (2001). Parents' hostility in dyadic marital and triadic family settings and children's behavior problems. *Journal of Consulting and Clinical Psychology, 69*, 727–734.
- Gottman, J. M. (1994). *What predicts divorce? The relationship between marital processes and marital outcomes*. Hillsdale, NJ: Erlbaum.
- Gottman, J. M. (1999). *The marriage clinic: A scientifically based marital therapy*. New York: Norton.
- Gottman, J. M., & Levenson, R. W. (1999a). How stable is marital interaction over time? *Family Process, 38*, 159–165.
- Gottman, J. M., & Levenson, R. W. (1999b). What predicts change in marital interaction over time? A study of alternative medicine. *Family Process, 38*, 143–158.
- Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the Difficulties in Emotion Regulation Scale. *Journal of Psychopathology and Behavioral Assessment, 26*, 41–54.
- Grotevant, H. D., & Carlson, C. I. (1989). *Family assessment: A guide to methods and measures*. New York: Guilford Press.
- Hahlweg, K. (2004). 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., Reisner, L., Kohli, G., Vollmer, M., Schindler, L., & Revenstorf, D. (1984). Development and validity of a new system to analyze interpersonal communication: Kategoriensystem für partnerschaftliche Interaktion (KPI) [Interactional Coding System (ICS)]. In K. Hahlweg & N. S. Jacobson (Eds.), *Marital interaction: Analysis and modification* (pp. 182–198). New York: Guilford Press.
- Hawkins, M. W., Carrère, S., & Gottman, J. M. (2002). Marital sentiment override: Does it influence couples' perceptions? *Journal of Marriage and Family, 64*, 193–201.
- Hayes, S. C., Nelson, R. O., & Jarrett, R. B. (1987). The treatment utility of assessment: A functional approach to evaluating assessment quality. *American Psychologist, 42*, 963–974.
- Haynes, S. N. (2001). Clinical applications of analogue behavioral observation: Dimensions of psychometric evaluation. *Psychological Assessment, 13*, 73–85.
- Haynes, S. N., Floyd, F. J., Lemsky, C., Rogers, E., Winemiller, D., Heilman, N., et al. (1992). The Marital Satisfaction Questionnaire for older persons. *Psychological Assessment, 4*, 473–482.
- Haynes, S. N., & Jensen, B. J. (1979). The interview as a behavioral assessment instrument. *Behavioral Assessment, 1*, 97–106.
- Haynes, S. N., Jensen, B., Wise, E., & Sherman, D. (1981). The marital intake interview: A multimethod criterion validity evaluation. *Journal of Consulting and Clinical Psychology, 49*, 379–387.
- Haynes, S. N., Leisen, M. B., & Blaine, D. D. (1997). Design of individualized behavioral treatment programs using functional analytic clinical case models. *Psychological Assessment, 9*, 334–348.
- Haynes, S. N., & O'Brien, W. H. (2000). *Principles and practice of behavioral assessment*. New York: Kluwer Academic.
- Haynes, S. N., Richard, D. R., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment, 7*, 238–247.
- Heavey, C. L., Layne, C., & Christensen, A. (1993). Gender and conflict structure in marital interaction: A replication and extension. *Journal of Consulting and Clinical Psychology, 61*, 16–27.
- Heim, S. C., & Snyder, D. K. (1991). Predicting depression from marital distress and attributional processes. *Journal of Marital and Family Therapy, 17*, 67–72.
- Heyman, R. E. (2001). Observation of couple conflicts: Clinical assessment applications, stubborn truths, and shaky foundations. *Psychological Assessment, 13*, 5–35.
- Heyman, R. E. (2004). Rapid Marital Interaction Coding System (RMICS). In P. K. Kerig & D. H. Baucom (Eds.), *Couple observational coding systems* (pp. 67–94). Mahwah, NJ: Erlbaum.
- Heyman, R. E., Feldbau-Kohn, S. R., Ehrensaft, M. K., Langhinrichsen-Rohling, J., & O'Leary, K. D. (2001). Can questionnaire reports correctly classify relationship distress and partner physical abuse? *Journal of Family Psychology, 15*, 334–346.

- Heyman, R. E., & Slep, A. M. S. (2004). Analogue behavioral observation. In M. Hersen, E. M. Heiby, & S. N. Haynes (Eds.), *Comprehensive handbook of psychological assessment: Vol. 3. Behavioral assessment* (pp. 162–180). New York: Wiley.
- Heyman, R. E., Weiss, R. L., & Eddy, J. M. (1995). Marital Interaction Coding System: Revision and empirical evaluation. *Behaviour Research and Therapy*, *33*, 737–746.
- Hooley, J. M., & Teasdale, J. D. (1989). Predictors of relapse in unipolar depressives: Expressed emotion, marital distress, and perceived criticism. *Journal of Abnormal Psychology*, *98*, 229–235.
- Hunsley, J., Best, M., Lefebvre, M., & Vito, D. (2001). The seven-item short form of the Dyadic Adjustment Scale: Further evidence for construct validity. *American Journal of Family Therapy*, *29*, 325–335.
- Jacob, T., & Tennenbaum, D. L. (1988). *Family assessment: Rationale, methods, and future directions*. New York: Plenum Press.
- Jones, A. C., & Chao, C. M. (1997). Racial, ethnic and cultural issues in couples therapy. In W. K. Halford & H. J. Markman (Eds.), *Clinical handbook of marriage and couples interventions* (pp. 157–176). New York: Wiley.
- Karpel, M. A. (1994). *Evaluating couples: A handbook for practitioners*. New York: Norton.
- Kasian, M., & Painter, S. L. (1992). Frequency and severity of psychological abuse in a dating population. *Journal of Interpersonal Violence*, *7*, 350–364.
- Kerig, P. K., & Baucom, D. H. (Eds.). (2004). *Couple observational coding systems*. Mahwah, NJ: Erlbaum.
- Kiecolt-Glaser, J. K., & Newton, T. L. (2001). Marriage and health: His and hers. *Psychological Bulletin*, *127*, 472–503.
- Kline, G. H., Julien, D., Baucom, B., Hartman, S., Gilbert, K., Gonzalez, T., & Markman, H. J. (2004). The Interactional Dimensions Coding System: A global system for couple interactions. In P. K. Kerig & D. H. Baucom (Eds.), *Couple observational coding systems* (pp. 113–126). Mahwah, NJ: Erlbaum.
- Kreider, R. M., & Fields, J. M. (2002). Number, timing, and duration of marriages and divorces: 1996. In *Current Population Reports* (No. P70–80). Washington, DC: U.S. Census Bureau.
- L'Abate, L. (1994). *Family evaluation: A psychological approach*. Thousand Oaks, CA: Sage.
- L'Abate, L., & Bagarozzi, D. A. (1993). *Sourcebook of marriage and family evaluation*. New York: Brunner/Mazel.
- Laumann, E. O., Gagnon, J. H., Michael, R. T., & Michaels, S. (1994). *The social organization of sexuality*. Chicago: University of Chicago Press.
- Law, D. D., & Crane, D. R. (2000). The influence of marital and family therapy on health care utilization in a health-maintenance organization. *Journal of Marital and Family Therapy*, *26*, 281–291.
- Leonard, K. E., & Senchak, M. (1996). Prospective prediction of husband marital aggression within newlywed couples. *Journal of Abnormal Psychology*, *105*, 369–380.
- Lin, E., Goering, P., Offord, D. R., Campbell, D., & Boyle, M. H. (1996). The use of mental health services in Ontario: Epidemiologic findings. *Canadian Journal of Psychiatry*, *41*, 572–577.
- Locke, H. J., & Wallace, K. M. (1959). Short marital adjustment and prediction tests: Their reliability and validity. *Marriage and Family Living*, *21*, 251–255.
- LoPiccolo, J., & Steger, J. C. (1974). The Sexual Interaction Inventory: A new instrument for assessment of sexual dysfunction. *Archives of Sexual Behavior*, *3*, 585–595.
- Lord, C. C. (1999). *Stability and change in interactional behavior in early marriage*. Unpublished doctoral dissertation, Stony Brook University, State University of New York.
- Malik, N. M., & Lindahl, K. M. (2004). System for Coding Interactions in Dyads (SCID). In P. K. Kerig & D. H. Baucom (Eds.), *Couple observational coding systems* (pp. 173–190). Mahwah, NJ: Erlbaum.
- Mash, E. J., & Foster, S. L. (2001). Exporting analogue behavioral observation from research to clinical practice: Useful or cost-defective? *Psychological Assessment*, *13*, 86–98.
- Mash, E. J., & Hunsley, J. (2004). Behavioral assessment: Sometimes you get what you need. In M. Hersen, E. M. Heiby, & S. N. Haynes (Eds.), *Comprehensive handbook of psychological assessment: Vol. 3. Behavioral assessment* (pp. 489–501). New York: Wiley.
- McConaghy, N. (1998). Assessment of sexual dysfunction and deviation. In A. S. Bellack & M. Hersen (Eds.), *Behavioral assessment: A practical handbook* (4th ed., pp. 315–341). Boston: Allyn & Bacon.
- Means-Christensen, A. J., Snyder, D. K., & Negy, C. (2003). Assessing nontraditional couples: Validity of the Marital Satisfaction Inventory—Revised (MSI-R) with gay, lesbian, and cohabiting heterosexual couples. *Journal of Marital and Family Therapy*, *29*, 69–83.
- Metz, M. E. (1993). *Manual for the Styles of Conflict Inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Morey, L. C. (1991). *Personality assessment inventory*. Lutz, FL: Psychological Assessment Resources.
- Mrazek, P. J., & Haggerty, R. J. (Eds.). (1994). *Reducing risks for mental disorders: Frontiers for preventive intervention research*. Washington, DC: National Academy Press.
- Northey, W. F., Jr. (2002). Characteristics and clinical practices of marriage and family therapists: A national survey. *Journal of Marital and Family Therapy*, *28*, 487–494.
- Norton, R. (1983). Measuring marital quality: A critical look at the dependent variable. *Journal of Marriage and the Family*, *45*, 141–151.
- O'Leary, K. D. (Ed.). (1987). *Assessment of marital discord: An integration for research and clinical practice*. Hillsdale, NJ: Erlbaum.
- O'Leary, K. D., & Maiuro, R. D. (2001). *Psychological abuse in violent domestic relations*. New York: Springer.
- Pasch, L. A., Bradbury, T. N., & Davila, J. (1997). Gender, negative affectivity, and observed social support behavior in marital interaction. *Personal Relationships*, *4*, 361–378.
- Pasch, L. A., Harris, K. W., Sullivan, K. T., & Bradbury, T. N. (2004). The Social Support Interaction Coding System (SSICS). In P. K. Kerig & D. H. Baucom (Eds.), *Couple observational coding systems* (pp. 319–334). Mahwah, NJ: Erlbaum.
- Persons, J. B. (1989). *Cognitive therapy in practice: A case formulation approach*. New York: Norton.
- Pretzer, J. L., Epstein, N., & Fleming, B. (1992). The Marital Attitude Survey: A measure of dysfunctional attitudes and expectancies. *Journal of Cognitive Psychotherapy*, *5*, 131–148.
- Rathus, J. H., & Feindler, E. L. (2004). *Assessment of partner violence: A handbook for researchers and practitioners*. Washington, DC: American Psychological Association.
- Sayers, S. L., & Sarwer, D. B. (1998). Assessment of marital dysfunction. In A. S. Bellack & M. Hersen (Eds.), *Behavioral assessment: A practical handbook* (4th ed., pp. 293–314). Boston: Allyn & Bacon.
- Schaefer, M. T., & Olson, D. H. (1981). Diagnosing intimacy: The PAIR inventory. *Journal of Marital and Family Therapy*, *7*, 47–60.
- Schumm, W. R., Paff-Bergen, L. A., Hatch, R. C., Obiorah, F. C., Copeland, J. M., Meens, L. D., & Bugaighis, M. A. (1986). Concurrent and discriminant validity of the Kansas Marital Satisfaction Scale. *Journal of Marriage and the Family*, *48*, 381–387.
- Shapiro, A. F., & Gottman, J. M. (2004). The Specific Affect Coding System (SPAFF). In P. K. Kerig & D. H. Baucom (Eds.), *Couple observational coding systems* (pp. 191–208). Mahwah, NJ: Erlbaum.
- Shaver, P. R., & Mikulincer, M. (2004). What do self-report attachment measures assess? In W. S. Rholes & J. A. Simpson (Eds.), *Adult attachment: Theory, research, and clinical implications* (pp. 17–54). New York: Guilford Press.
- Sher, T. G., Baucom, D. H., & Larus, J. M. (1990). Communication patterns and response to treatment among depressed and nondepressed maritally distressed couples. *Journal of Family Psychology*, *4*, 63–79.
- Shumway, S. T., Wampler, R. S., Dersch, C., & Arredondo, R. (2004). A

- place for marriage and family services in employee assistance programs (EAPs): A survey of EAP client problems and needs. *Journal of Marital and Family Therapy*, 30, 71–79.
- 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, 479–502.
- Snyder, D. K. (1997). *Manual for the Marital Satisfaction Inventory—Revised*. Los Angeles: Western Psychological Services.
- Snyder, D. K., & Abbott, B. V. (2002). Couple distress. In M. M. Antony & D. H. Barlow (Eds.), *Handbook of assessment and treatment planning for psychological disorders* (pp. 341–374). New York: Guilford Press.
- Snyder, D. K., & Aikman, G. G. (1999). The Marital Satisfaction Inventory—Revised. In M. E. Maruish (Ed.), *Use of psychological testing for treatment planning and outcomes assessment* (2nd ed., pp. 1173–1210). Mahwah, NJ: Erlbaum.
- Snyder, D. K., Cavell, T. A., Heffer, R. W., & Mangrum, L. F. (1995). Marital and family assessment: A multifaceted, multilevel approach. In R. H. Mikesell, D. D. Lusterman, & S. H. McDaniel (Eds.), *Integrating family therapy: Handbook of family psychology and systems theory* (pp. 163–182). Washington, DC: American Psychological Association.
- Snyder, D. K., Cepeda-Benito, A., Abbott, B. V., Gleaves, D. H., Negy, C., Hahlweg, K., & Laurenceau, J. P. (2004). Cross-cultural applications of the Marital Satisfaction Inventory—Revised (MSI-R). In M. E. Maruish (Ed.), *Use of psychological testing for treatment planning and outcomes assessment* (3rd ed., pp. 603–623). Mahwah, NJ: Erlbaum.
- Snyder, D. K., Mangrum, L. F., & Wills, R. M. (1993). Predicting couples' response to marital therapy: A comparison of short- and long-term predictors. *Journal of Consulting and Clinical Psychology*, 61, 61–69.
- Snyder, D. K., & Rice, J. L. (1996). Methodological issues and strategies in scale development. In D. H. Sprenkle & S. M. Moon (Eds.), *Research methods in family therapy* (pp. 216–237). New York: Guilford Press.
- Snyder, D. K., & Whisman, M. A. (Eds.). (2003). *Treating difficult couples: Helping clients with coexisting mental and relationship disorders*. New York: Guilford Press.
- 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.
- Straus, M. A. (1979). Measuring intrafamily conflict and violence: The Conflict Tactics (CT) Scales. *Journal of Marriage and the Family*, 41, 75–88.
- Straus, M. A., Hamby, S. L., Boney-McCoy, S., & Sugarman, D. B. (1996). The revised Conflict Tactics Scales (CTS2): Development and preliminary psychometric data. *Journal of Family Issues*, 17, 283–316.
- Strosahl, K. D., & Robinson, P. J. (2004). Behavioral assessment in the era of managed care: Understanding the present, preparing for the future. In M. Hersen, E. M. Heiby, & S. N. Haynes (Eds.), *Comprehensive handbook of psychological assessment: Vol. 3. Behavioral assessment*. New York: Wiley.
- Suhr, J. A., Cutrona, C. E., Krebs, K. K., & Jensen, S. L. (2004). The Social Support Behavior Code (SSBC). In P. K. Kerig & D. H. Baucom (Eds.), *Couple observational coding systems* (pp. 311–318). Mahwah, NJ: Erlbaum.
- Swindle, R., Heller, K., Pescosolido, B., & Kikuzawa, S. (2000). Responses to nervous breakdowns in America over a 40-year period: Mental health policy implications. *American Psychologist*, 55, 740–749.
- Tanaka-Matsumi, J. (2004). Individual differences and behavioral assessment. In S. N. Haynes & E. M. Heiby (Eds.), *Behavioral assessment*. Hoboken, NJ: Wiley.
- Tolman, R. M. (1989). The development of a measure of psychological maltreatment of women by their male partners. *Violence and Victims*, 4, 159–177.
- Touliatos, J., Perlmutter, B. F., Straus, M. A., & Holden, G. W. (Eds.). (2001). *Handbook of family measurement techniques* (Vols. 1–3). Thousand Oaks, CA: Sage.
- Veroff, J., Kulka, R. A., & Douvan, E. (1981). *Mental health in America: Patterns of help seeking from 1957 to 1976*. New York: Basic Books.
- Vivian, D., Langhinrichsen-Rohling, J., & Heyman, R. E. (2004). Thematic Coding of Dyadic Interaction. In P. K. Kerig & D. H. Baucom (Eds.), *Couple observational coding systems* (pp. 273–288). Mahwah, NJ: Erlbaum.
- Weiss, R. L. (1980). Strategic behavioral marital therapy: Toward a model for assessment and intervention. In J. P. Vincent (Ed.), *Advances in family intervention, assessment, and theory* (Vol. 1, pp. 229–271). Greenwich, CT: JAI Press.
- Weiss, R. L., & Birchler, G. R. (1975). *Areas of change questionnaire*. Unpublished manuscript, University of Oregon.
- Weiss, R. L., & Heyman, R. E. (1997). A clinical-research overview of couples interactions. In W. K. Halford & H. J. Markman (Eds.), *Clinical handbook of marriage and couples intervention* (pp. 13–41). New York: Wiley.
- Weiss, R. L., & Perry, B. A. (1983). The Spouse Observation Checklist: Developments and clinical applications. In E. E. Filsinger (Ed.), *Marriage and family assessment: A sourcebook for family therapy* (pp. 65–84). Beverly Hills, CA: Sage.
- Whisman, M. A. (1999). Marital dissatisfaction and psychiatric disorders: Results from the National Comorbidity Survey. *Journal of Abnormal Psychology*, 108, 701–706.
- Whisman, M. A. (2001). Marital adjustment and outcome following treatments for depression. *Journal of Consulting and Clinical Psychology*, 69, 125–129.
- Whisman, M. A., & Bruce, M. L. (1999). Marital dissatisfaction and incidence of major depressive episode in a community sample. *Journal of Abnormal Psychology*, 108, 674–678.
- Whisman, M. A., Dixon, A. E., & Johnson, B. (1997). Therapists' perspectives of couple problems and treatment issues in couple therapy. *Journal of Family Psychology*, 11, 361–366.
- Whisman, M. A., Sheldon, C. T., & Goering, P. (2000). Psychiatric disorders and dissatisfaction with social relationships: Does type of relationship matter? *Journal of Abnormal Psychology*, 109, 803–808.
- Whisman, M. A., & Uebelacker, L. A. (2003). Comorbidity of relationship distress and mental and physical health problems. In D. K. Snyder & M. A. Whisman (Eds.), *Treating difficult couples: Helping clients with coexisting mental and relationship disorders* (pp. 3–26). New York: Guilford Press.
- Wieder, G. B., & Weiss, R. L. (1980). Generalizability theory and the coding of marital interactions. *Journal of Consulting and Clinical Psychology*, 48, 469–477.

Received February 5, 2004

Revision received August 11, 2004

Accepted May 16, 2005 ■