Emotional Behavior in Long-Term Marriage

Laura L. Carstensen
Department of Psychology Stanford University

John M. Gottman
Department of Psychology University of Washington

Robert W. Levenson
Department of Psychology University of California, Berkeley

ABSTRACT

In exploring the emotional climate of long-term marriages, this study used an observational coding system to identify specific emotional behaviors expressed by middle-aged and older spouses during discussions of a marital problem. One hundred and fifty-six couples differing in age and marital satisfaction were studied. Emotional behaviors expressed by couples differed as a function of age, gender, and marital satisfaction. In older couples, the resolution of conflict was less emotionally negative and more affectionate than in middle-aged marriages. Differences between husbands and wives and between happy and unhappy marriages were also found. Wives were more affectively negative than husbands, whereas husbands were more defensive than wives, and unhappy marriages involved greater exchange of negative affect than happy marriages.

We wish to acknowledge National Institute on Aging Grants AG08816 and AG07476 and Research Scientist Award 1K02MH00257.

Correspondence may be addressed to Laura L. Carstensen, Department of Psychology, Stanford University, Stanford, California, 94305.
Correspondence may be addressed to Robert W. Levenson, Department of Psychology, University of California, 3210 Tolman Hall, Berkeley, California, 94720.

Received: May 8, 1992
Revised: June 24, 1994
Accepted: June 24, 1994

We had two primary goals in this research. The first was to explore differences in emotional expression in intimate relationships in middle-aged and older marriages. The second was to test the generalizability to middle-aged and older couples of findings from earlier research investigating the influence of gender and marital satisfaction on emotional interaction in young married couples. The study was based on the analysis of videotaped interactions of happy and unhappy couples discussing conflicts in their relationships (see also Levenson, Carstensen, & Gottman, 1993). By pursuing our research questions in both happy and unhappy marriages, we were able to examine the pervasiveness of age and gender differences across these very different types of couples. In the following sections, we elaborate our rationale and hypotheses for the research.

Emotion and Aging
Early theories of old age depicted emotion as dampened, rigid, and flat (Banham, 1951; Looft, 1972). Yet, a growing body of empirical evidence is painting quite a different picture. Findings from several laboratories suggest age-related improvement in the control of emotion (Lawton, Kleban, Rajagopal, & Dean, 1992) and emotional understanding (Labouvie-Vief & De Voe, 1991; Labouvie-Vief, Hakim-Larson, De Voe, & Schoeberlein, 1989). Self-reported intensity of emotion is comparable for old and young people (Levenson, Carstensen, Friesen, & Ekman, 1991; Malatesta & Kalnok, 1984), and some research suggests that emotion becomes increasingly salient with age (Carstensen, 1992; Carstensen & Turk-Charles, 1994; Fredrickson & Carstensen, 1990; Hashtroudi, Johnson, & Chrosniak, 1990). This evidence, in tandem with findings that older people experience relatively low rates of psychological disorders (George, Blazer, Winfield-Laird, Leaf, & Fischback, 1988), suggests that emotion may represent a life domain in which functioning is well preserved or even improves with age.

If improved functioning does indeed characterize emotion in old age, it is important to undertake examination of the mechanisms underlying such change. Theoretical arguments have been advanced that older people proactively design their environments in the service of emotion regulation (Carstensen, 1993; Lawton, 1989; Schulz, 1985). In socioemotional selectivity theory, for example, Carstensen (1991, 1993) argued that because emotion becomes more salient with age, people mold social interactions to maximize emotional benefits. However, we know of almost no studies of older adults that have examined emotion expressed within a social context. If older people do "manage" social interactions toward the optimization of positive affect, one should see evidence of this in the interactional dynamics surrounding discussion of emotionally charged topics.

The marital relationship offers an ideal context in which to examine emotion. First, emotions experienced in the context of marriage run the gamut from joy to contempt, and the negotiation of affectively laden issues is required in the best of marriages. Second, a growing body of research, based on cross-sectional and longitudinal findings, suggests that—after a decline in marital satisfaction in midlife—marriages become increasingly positive as the partners enter old age (Brubaker, 1990; Guilford & Bengtson, 1979; Levenson et al., 1993). Thus, long-term marriages theoretically represent social relationships that involve both positive and negative affect and show age-related change in a positive direction. By examining the expression of specific emotions, emotional exchanges, and emotional sequences in discussions of marital conflicts, one should be able to identify age-related differences in the strategies used in negotiating emotional interactions.

**Gender Differences in Marital Interactions**

Our earlier research (Gottman & Levenson, 1988; Levenson, Carstensen, & Gottman, 1994; Levenson & Gottman, 1983), as well as research from other laboratories (Christensen & Heavey, 1990; Notarius & Johnson, 1982; Schaap, 1982), suggests that there are reliable gender differences in marital behavior. Women are more confronting and more expressive emotionally and both more negative and more positive than men. Men are less expressive emotionally and are, in the face of conflict, more likely than women to be defensive and withdrawn and to engage in "stonewalling" behavior (i.e., nonexpressiveness when experiencing strong negative emotion).

However, these findings are based almost exclusively on research with relatively young couples. Virtually no research has addressed whether or not similar interactional patterns are present in older married couples. It is possible that the behavioral patterns observed in younger couples characterize only new marriages. Over time, gender-linked patterns may be overshadowed by idiosyncratic patterns of individual couples. It is also possible that gender differences diminish with age. Several researchers (e.g., Gutmann, 1987; Hyde & Phillis, 1979; Keith & Brubaker, 1979) have suggested that men and women become more similar as they age. If such convergence occurs in the realm of emotional
expression, then the oft-reported finding that wives are more likely to engage in conflict and husbands are more likely to withdraw might not hold in older marriages.

**Overview of the Use of Observational Methods in Research on Marriage and Marital Satisfaction**

The application of observational methods to research on marriage began in the early 1970s, addressing the fundamental question of how satisfied and dissatisfied couples differed in their resolution of conflict. In more than 50 studies conducted in the United States and cross nationally during the next two decades, observational methodology has revealed consistent differences between satisfied and dissatisfied couples (for reviews, see Gottman, 1979; Markman & Notarius, 1987; Schaap, 1982, 1984; Schaap, Buunk, & Kerkstra, 1988; Weiss & Heyman, 1990; Weiss & Summers, 1983).

Previous observational research has produced a number of consistent results. First, negative emotional behavior, such as expressed anger, sadness, contempt, and other negative emotions, appears to be the best discriminator between satisfied and dissatisfied marriages (see Schaap et al., 1988, for a review). Second, sequences of emotion expressed during marital interaction differentiate happy from unhappy couples. When the expression of negative affect by one spouse is reciprocated by negative affect from the other spouse, a pattern termed *negative affect reciprocity*, couples are likely to be dissatisfied. The temporal sequential stability or continuance of negative affective states across time also characterizes dissatisfied marriages. Third, differences between satisfied and dissatisfied couples are most pronounced when couples engage in attempts to resolve conflict (Gottman, 1979).

How representative of marriage were these observational studies, and how generalizable were their conclusions? Krokoff (1987) pointed out that the mean age of couples in past marital interaction studies was 29.94 years, whereas the mean age of the U.S. population is 44.70 years (and rising). Hence, these observational studies were primarily studies of relatively young couples in relatively new marriages. Their generality to middle-aged and older couples remains an open question.

We are aware of only two observational studies that have been conducted with older couples. One was an unpublished dissertation by Illig (1977), and the other was a published study by Zietlow and Sillars (1988). Zietlow and Sillars studied 59 couples during marital discussions. They found that "aging couples are characteristically happy, although inexpressive by modern standards" (p. 241). The emotional quality of the discussions of the older couples they studied differed as a function of the salience of the topic. When the couples discussed nonsalient topics, they were quite cordial; when they discussed salient topics, they were quite confronting. Both Illig (1977) and Zietlow and Sillars (1988) noted that the interactions of older couples had more temporal sequential predictability than the interactions of younger couples, particularly greater negative reciprocity.

These studies represent the first use of observational methods with older couples, but they are not without problems. Zietlow and Sillars (1988) recruited all of their couples from one Lutheran church in Milwaukee, which probably limits the generality of the results. Neither their study nor Illig's (1977) investigation assessed or controlled for marital satisfaction. Given the real possibility that couples in their age groups differed in marital satisfaction, it is impossible to know whether the reported findings were characteristic of different age groups, of happy or unhappy marriages, or of both.

**The Present Study**

Our aim in the present study was to explore affect expression and reciprocity in long-term marriages using a detailed observational system that allowed for the coding of specific emotions. First, we asked
whether patterns of affect and affect reciprocity were more positive in older than in middle-aged couples. We hypothesized that older couples would express more positive emotions and fewer negative emotions during conflict discussions than would middle-aged couples. Second, we asked whether relationships that negative affect and negative affect reciprocity have been found to have with marital unhappiness in previous research with younger marriages would also be found in middle-aged and older marriages. If negative affective expression diminishes or loses its aversive quality over the years, then it might not be as strongly related to marital dissatisfaction. Finally, we asked whether gender differences observed in young marriages would still be obtained in later-life marriages.

Method

Participants Constructing the experimental sample.

A three-stage sampling procedure was used to recruit couples from communities surrounding Berkeley, California (see Levenson et al., 1993, for a complete description of the procedure). After establishing characteristics of older couples in the local community in terms of marital satisfaction, ethnicity, and socioeconomic status through a random telephone survey, we established recruitment goals (e.g., a certain percentage of old, unhappy, White, and blue-collar couples) that would represent the population of interest. 1 A large sample of participants (N = 960) was then recruited by means of newspaper, radio, bulletin boards, and placards on city buses. These participants were screened by telephone and questionnaire responses to determine their suitability for the experiment. Screened participants were selected on the basis of recruitment goals for participation in the laboratory experiment.

Recruitment eligibility included length of marriage, age, and marital satisfaction criteria. To ensure that all couples were in long-term marriages, we required that middle-aged couples be married for at least 15 years, with the older spouse between 40 and 50 years of age. Older couples had to be married for at least 35 years, with the older spouse between 60 and 70 years old. 2 To ensure representation of both relatively happy and relatively unhappy couples in the sample, we recruited comparable numbers of happy and unhappy couples in both middle-aged and older subsamples. In addition, we required that the difference between spouses' marital satisfaction scores be no more than 20 points so that the spouses in a couple would not differ substantially in their feelings about the marriage. Because older couples were determined to be happier, on average, than middle-aged couples during telephone screening, we initially used different marital satisfaction cutoffs in recruiting the two subsamples. However (as described later), so as not to confound marital satisfaction with age, we conducted all data analyses using satisfaction groupings based on the median satisfaction score for the entire sample.

Additional criteria for participation were as follows: (a) Neither spouse had retired; (b) English was the native language or language customarily spoken at home; (c) the age difference between spouses was no more than 5 years; and (d) neither spouse was alcoholic (as indicated by scoring below 7 on the Michigan Alcoholism Screening Test; Selzer, 1971). The rationale for these additional satisfaction and age criteria was to make our sample representative of the modal long-term marriage, in which couples are relatively close in marital satisfaction and age, and to avoid having spouses fall into different satisfaction or age groups.

The final laboratory sample consisted of 156 couples assigned to one of four groups according to age and marital satisfaction using the grand average of both spouses' scores on two measures of marital satisfaction, the Locke—Williamson measure (Burgess, Locke, & Thomes, 1971) and the Locke—Wallace measure (Locke & Wallace, 1959). For the entire sample, the mean for this aggregated score was 111.5 (SD = 17.0) and the median was 115.1 (range = 42.8—160.0). This median score was used to make the final assignments to experimental groups as follows: (a) middle-aged satisfied (n = 35), (b)
middle-aged dissatisfied \((n = 47)\), (c) old satisfied \((n = 43)\), and (d) old dissatisfied \((n = 31)\).

**Sample demographics.**

For middle-aged couples, the mean ages were 44.3 \( (SD = 2.9) \) for husbands and 43.3 \( (SD = 2.9) \) for wives. For older couples, the mean ages were 63.6 \( (SD = 2.9) \) for husbands and 62.2 \( (SD = 3.2) \) for wives. For the middle-aged satisfied group, the mean aggregated satisfaction score was 122.7 \( (SD = 5.8) \) and the mean duration of marriage in years was 21.3 \( (SD = 3.4) \); for the middle-aged dissatisfied group, 98.7 \( (SD = 13.7) \) and 20.9 \( (SD = 3.4) \); for the old satisfied group, 125.1 \( (SD = 8.5) \) and 40.4 \( (SD = 3.9) \); and, for the old dissatisfied group, 99.3 \( (SD = 15.3) \) and 39.8 \( (SD = 3.1) \).

The resulting sample was 85.9% White, 5.8% African American, 2.6% Hispanic, 2.6% Asian, and 2.2% with spouses of different races. Experimental groups were recruited such that they matched the general demographic balance of the geographic area (as determined by the random telephone survey) for older couples in terms of religion, ethnicity, and socioeconomic status (Levenson et al., 1993). Also, the groups themselves were matched for religion, ethnicity, and socioeconomic status.

**Confounding of age and marital duration.**

One particular feature of our sampling strategy requires separate mention, namely, our decision to focus on long-term first marriages in these two age groups. We focused our investigation on these kinds of marriages because they are most representative of marriages in the current middle-aged and older cohorts and because we were particularly interested in learning about couples who had married in early adulthood and had managed to stay together ever since. One implication of this decision was that we would be confounding age with duration of marriage.

It is important to recognize that, in this regard, our sample is a reasonable proxy for marriages in the real world, where age and length of marriage are also typically confounded. Of course, it would have been possible to separate age and marital duration, but doing so would have meant recruiting a sample of couples who would be very unrepresentative of marriages in at least one of the age cohorts. For example, had we required that both older and middle-aged couples be married for 15 years, the resulting sample would have been composed of middle-aged couples who married in their 20s and 30s and older couples who married in their 40s and 50s. Thus, age at the time of marriage would have been confounded with current age. Furthermore, unless we recruited older couples who married for the first time in their 40s and 50s (a highly unrepresentative group), we would probably be studying remarriages in our older sample versus first marriages in our middle-aged sample, another serious confound.

In the final analysis, we chose to study the largest sample we thought we could manage (a sample twice the size of our largest previous observational study) and included in this sample couples that we believed were most representative of marriages in these age cohorts. Thus, our modal couple was in a first marriage that had lasted for most of their adult lives. For us, the advantages of this sampling strategy outweigh the disadvantages. However, we ask readers to be aware that, when we mention age differences, these may result from differences in the duration of marriage as well.

**Procedure**

The procedures used in this experiment were modeled after those developed by Levenson and Gottman (1983). Couples came to the laboratory after having not spoken to each other for at least 8 hr. Recording devices for obtaining physiological measures were attached, and couples engaged in three conversational interactions: (a) discussing the events of the day, (b) discussing a problem area of
continuing disagreement in their marriage, and (c) discussing a mutually agreed on pleasant topic. Before the problem area interaction, couples completed the Couple's Problem Inventory (Gottman, Markman, & Notarius, 1977), in which they rated the perceived severity of 10 marital issues on a scale ranging from 0 to 100. Using these ratings, the experimenter helped couples pick a topic that both spouses had rated highly, interviewed them about the topic, and helped them focus on the key area of disagreement. In our experience, this interview serves to make the ensuing discussion more personal and less abstract.

Each conversation lasted for 15 min, preceded by a 5-min silent period. A split-screen video recording was made of the entire session. Couples were paid $150 for participating in the laboratory experiment, which consisted of the interaction session and two subsequent sessions in which spouses separately viewed the video recording and provided ratings of how they were feeling during the interaction. Although physiological measures and self-ratings of interactions were obtained (see Levenson et al., 1994), we focus here only on the behavioral coding of affect that occurred during the problem area conversation.

Apparatus

Two remotely controlled high-resolution video cameras that were partially concealed behind darkened glass were used to obtain frontal views of each spouse's face and upper torso. These images were combined into a single split-screen image with a video special effects generator and were recorded on a VHS videocassette recorder. Two lavaliere microphones were used to record the spouses' conversations.

Data Reduction

A team of coders using the Specific Affect Coding System (SPAFF; Gottman & Krokoff, 1989; SPAFF Version 2.0, Gottman, 1989), which dismantles affect into specific positive and negative emotions and includes codes for both speakers and listeners, coded the problem area discussions. SPAFF is a cultural informant system in which coders, working with videotapes, consider a gestalt consisting of verbal content, voice tone, context, facial expression, gestures, and body movement. SPAFF treats the stream of behavior as continuous (rather than segmenting it into time blocks or turns at speech), and, thus, codes can be given at any time. The code best describing the affect of a spouse is indicated on a computerized dial until a change in behavior occurs such that another code better reflects the affective state of the spouse.

For speakers, the positive affect codes are interest, affection, humor, validation (i.e., acknowledgment of partner's feelings), and joy. The negative affect codes are anger, contempt, disgust, belligerence, domineering, defensiveness, fear/tension/worry, sadness, and whining. There is also a neutral speaker code. For listeners, codes are positive, negative, and neutral. These codes are based on the facial expressions of the listener. There is an additional code for listener disengagement called stonewalling.

Reliability for SPAFF codes is based on the second-by-second concordance of observers throughout the 15-min (900-s) interaction period. When more than one code appears during a 1-s period, reliability is based on the code indicated for the longer period of time in milliseconds. The statistic used to calculate reliability was Cohen's kappa, which controls for agreement by chance alone and provides a single reliability index for the entire coding system (Bakeman & Gottman, 1986). For this study, the overall mean kappa was 0.64, and the mean z score (kappa divided by the standard deviation of kappa) was 19.25 (the mean z score must exceed 1.96 for agreement to be significantly greater than chance). The mean kappa for listener codes was 0.71, and the z score was 16.92 (p < .001). For speaker codes, the kappa was 0.60 and the z score was 15.02 (p < .001). To control for differences among couples in the
total number of codes given, we converted SPAFF data for any given spouse in any given couple to percentages of the total number of speaker and listener codes expressed by that spouse and his or her partner (i.e., the sum of the 15 speaker codes and 4 listener codes for both spouses).

**Results**

**Overview**

SPAFF codes were examined at both specific and relatively global levels. First, to describe differences in the profiles of affect by age, marital satisfaction, and gender while preserving the detail provided by the 15 specific affect speaker codes and 4 listener codes, we entered the SPAFF codes into an overall multivariate analysis of variance (MANOVA). Second, to test a small number of conceptually interesting hypotheses about affect sequences, we collapsed the specific SPAFF speaker and listener codes into 3 global codes (positive, negative, and neutral) and submitted these codes to sequential analysis. (In the findings discussed subsequently, the .05 rejection level was used unless otherwise noted.)

**Specific Affect**

SPAFF codes were analyzed by means of 2 (age: middle-aged vs. older couples) × 2 (satisfaction: happy vs. unhappy) × 2 (spouse: husband vs. wife) MANOVAs and analyses of variance (ANOVAs); spouse was treated as a repeated measure. An overall MANOVA was computed for the 14 nonneutral speaker codes and the 3 nonneutral listener codes (because the data were expressed as percentages of codes, for any given spouse, of the total number of codes expressed by both spouses, the neutral code was omitted to avoid linear dependency in the analysis). This was followed by univariate ANOVAs for the 15 speaker and 4 listener codes and for 3 global speaker codes: total speaker positive (sum of percentages of interest, affection, humor, validation, and joy), total speaker negative (sum of percentages of anger, contempt, disgust, belligerence, domineering, defensiveness, fear/tension/worry, sadness, and whining), and total speaker emotion (sum of percentages of positive and negative speaker codes).

The overall MANOVA revealed significant main effects of age, $F(17, 136) = 4.25, p < .001$; satisfaction, $F(17, 136) = 2.13, p = .009$; and spouse, $F(17, 136) = 5.37, p < .001$. None of the interactions in the overall MANOVA were significant. Means and $F$ values for the univariate age, satisfaction, and spouse main effects are presented in Table 1.

**Age effects.**

In terms of speaker behaviors, elderly couples were coded as being more affectionate than middle-aged couples. Middle-aged couples were coded as displaying more interest, humor, anger, disgust, belligerence, and whining than older couples. There were no age differences in listener behaviors.

Because we had previously found that older couples rated their marital problems as less severe than those of middle-aged couples (Levenson et al., 1993), we thought it important to establish that age differences found in emotional behavior did not merely reflect age differences in problem severity. Thus, for those behaviors for which significant age differences had been found, we reran our analyses using the average rated severity of marital problems as a covariate (these ratings had been obtained before the problem area discussion by means of the Couple's Problem Inventory [Gottman et al., 1977]).

Age-related differences in affection, $F(1, 146) = 6.45, p = .012$; interest, $F(1, 146) = 10.41, p = .002$;
humor, $F(1, 146) = 8.42$, $p = .004$; disgust, $F(1, 146) = 5.54$, $p = .020$; and belligerence, $F(1, 146) = 10.37$, $p = .002$, survived this test. The age difference in whining approached significance, $F(1, 146) = 3.74$, $p = .055$, whereas the age difference in anger was no longer significant, $F(1, 146) = 0.87$.

**Gender effects.**

In terms of speaker behaviors, wives were coded as showing more total emotion, negative emotion, anger, joy, contempt, whining, and sadness. Husbands were coded as being more defensive than wives. In terms of listener behavior, husbands were coded as being more neutral listeners than wives; wives were coded as being both more positive and more negative listeners than husbands.

Significant Satisfaction $\times$ Spouse interaction effects were found for several affect codes. In general, for these codes, wives were more emotional than husbands in unhappy marriages but not in happy marriages. This was true for total speaker emotion $F = 4.93$, $p < .03$; negative speaker emotion, $F(1, 152) = 4.02$, $p < .04$; contempt, $F(1, 152) = 6.42$, $p < .04$; and negative listener emotion, $F(1, 152) = 6.86$, $p < .01$. The mean percentages of total speaker emotion were 20.32 for unhappy wives and 17.21 for unhappy husbands, $t(152) = 3.05$, $p < .003$; they were 16.00 for happy wives and 16.74 for happy husbands. The mean percentages of negative speaker emotion were 16.41 for unhappy wives and 13.11 for unhappy husbands, $t(152) = 2.49$, $p < .01$; these percentages were 10.22 for happy wives and 11.29 for happy husbands. The mean percentages of contempt were 1.89 for unhappy wives and 0.97 for unhappy husbands, $t(152) = 3.95$, $p < .0001$; these percentages were 0.87 for happy wives and 0.64 for happy husbands. The mean percentages of listener emotion were 8.27 for unhappy wives and 6.46 for unhappy husbands, $t(152) = 2.09$, $p < .04$; they were 4.91 for happy wives and 5.10 for happy husbands.

Both a two-way Satisfaction $\times$ Spouse effects $F = 5.16$, $p < .02$, and a three-way Age $\times$ Satisfaction $\times$ Spouse interaction effect, $F = 4.58$, $p < .03$, were found for the neutral speaker code. However, follow-up $t$ tests revealed no differences between spouses in the various types of couples.

**Satisfaction effects.**

In terms of speaker behavior, happy couples were coded as showing more positive emotion and as being more neutral, more humorous, more affectionate, and more validating than unhappy couples. Unhappy couples were coded as showing more emotion overall and more negative emotion and as being more angry, contemptuous, sad, domineering, and belligerent than happily married couples. In terms of listener behavior, happy couples were coded as being more neutral and more positive listeners than unhappy couples; unhappy couples were coded as expressing more negative listening behavior than happy couples.

**Affect Sequences.**

In attempts to understand marital interaction, assessing the occurrence of single episodes of affect often does not tell the entire story. Rather, when one spouse expresses a certain type of affect, the type of affect expressed by the other spouse in response is often very informative. Unfortunately, with a coding system such as SPAFF, which has a large number of specific affect codes, the number of possible specific affect sequences is very large [ $2^2N^2$ for $N$ codes given to each spouse, i.e., 1,444 potential sequences]. To control for Type I error, we decided to limit the number of sequential analyses conducted by collapsing the 15 specific speaker codes and 4 listener codes into 3 global affect codes (positive, negative, and neutral). On an a priori basis, we selected five theoretically interesting sequences of global affect to analyze: (a) negative start-up (neutral affect by one spouse followed by negative affect by the
other spouse; Patterson, 1982), (b) negative continuance (negative affect by one spouse followed by negative affect by the other spouse; Gottman, 1979; Patterson, 1982), (c) deescalation (negative affect by one spouse followed by neutral affect by the other spouse), (d) positive continuance (positive affect by one spouse followed by positive affect by the other spouse; Gottman, 1979), and (e) neutral continuance (neutral affect by one spouse followed by neutral affect by the other spouse).

The sequential analyses proceeded in three steps. First, the SPAFF speaker and listener codes were collapsed into positive, negative, and neutral affect categories. Second, we computed \( z \)-score estimates of the strength of the sequential connection of each of the five sequences as initiated by each spouse. Third, we conducted 2 (age: middle-aged vs. older couples) × 2 (satisfaction: happy vs. unhappy) × 2 (spouse initiating the sequence: husband vs. wife) ANOVAs using the \( z \) scores for the five sequences as initiated by each spouse; we treated spouse as a repeated measure. Means and \( F \) values for the age, satisfaction, and spouse main effects are presented in Table 2.

Age effects.

There were no main effects by age and length of marriage in affect sequences. However, there was a significant Age × Satisfaction interaction for negative start-up. Decomposing this interaction revealed that older unhappy couples were less likely to manifest negative start-up sequences than any of the other groups, \( F(1, 152) = 4.62, p = .031 \). The mean ( \( z \)) negative start-ups were as follows: older unhappy couples, 0.11, and older happy couples, 0.85, \( t(152) = -2.70, p = .008 \); middle-aged unhappy couples, 0.74 (vs. older unhappy couples), \( t(152) = -2.30, p = .022 \); and middle-aged happy couples, 0.65 (vs. older unhappy couples), \( t(152) = -1.95, p = .05 \).

Gender effects.

Wives engaged in more positive continuance than husbands. A significant Satisfaction × Spouse interaction for deescalation revealed that husbands were more likely to deescalate than wives in unhappy marriages, but there were no spousal differences in happy marriages, \( F(1, 151) = 9.26, p = .022 \). Mean ( \( z \)) deescalations were 1.34 for unhappy husbands and 0.87 for unhappy wives, \( t(151) = 1.96, p = .048 \); means were 0.78 for happy husbands and 1.01 for happy wives, \( t(151) = -0.61, ns \).

Satisfaction effects.

Unhappy couples engaged in more negative continuance than happy couples.

Discussion

Findings from this project provide evidence for both stability and change in the nature of the marital relationship as it unfolds in later adulthood. Using a microanalytic emotional coding system, we documented considerable variation across couples in the ways that they discussed and attempted to resolve a problematic area in their relationship. In some ways, the patterns we observed in these middle-aged and older couples were highly similar to those seen in younger couples. In particular, interactional patterns associated with gender and level of marital satisfaction were evident even after many years of marriage. Nonetheless, we observed important differences in the realm of emotional expression related to the ages of the spouses and the duration of their marriages. Older couples, who in our study had also been married longer, expressed less negativity and more affection than middle-aged couples, even after differences in the severity of marital problems had been controlled. Thus, there is evidence suggesting an age-related positive affective trend within this highly intimate social relationship. In the following sections, we discuss in greater detail specific results by age, satisfaction, and gender.
Age

In this research, we used an experimental paradigm designed to elicit emotional interchange and studied the discussion of a marital problem, which is likely to generate negative affect. In a number of ways, the interactions of middle-aged couples were more emotional than those of older couples. We remind readers that we cannot disentangle age from length of marriage, and either could account for the findings. In terms of specific negative affect, middle-aged couples displayed higher levels of anger, disgust, belligerence, and whining than older couples. Possibly reflecting the attention-grabbing properties of negative affect (Hansen & Hansen, 1988; Pratto & John, 1991), the level of expressed interest was also higher in middle-aged couples. In terms of specific positive affect, middle-aged couples also displayed more humor than older couples.

For older couples, lower levels of negative affect could be seen as reflecting reduced engagement in the task or a general lessening of emotional intensity, or they could simply reflect the fact that older couples have less severe marital problems. However, we reject these interpretations on the basis of three findings. First, neither overall levels of emotion nor specific levels of tension, domineeringness, or contempt differentiated middle-aged and older couples, arguing against the notion that older people were simply not involved in the task. Second, the older couples expressed more of one kind of emotion than middle-aged couples: affection. This greater expression of affection in older couples argues against an overall dampening of emotional vitality. Third, even when age differences in the severity of marital problems were statistically controlled, old couples were less emotionally negative and more affectionate than middle-aged couples.

These results are consistent with the notion that older people actively use strategies that limit the experience of negative affect. Recall that all couples were discussing a conflict in their relationship. However, in doing so, older couples managed to express higher levels of affection and lower levels of negative affect toward their partner. Negative emotions were certainly expressed, but they were interwoven with affection. Thus, even when discussing areas of substantial disagreement with considerable potential for generating negative affect, elderly couples were more likely to communicate their positive feelings for each other.

Our analysis of affective sequences revealed one other age-related difference. Older unhappy couples were less likely to engage in sequences in which one spouse's neutral affect was followed by the other spouse's negative affect (i.e., negative start-up) than were any of the other three experimental groups. One of our overarching goals in undertaking this research has been to learn more about couples who have stayed together for long periods of time but are not satisfied with their marriages. This finding suggests that these couples may have learned to "leave well enough alone" by staying in affectively neutral interactive sequences and avoiding escalation to negative affect.

Collectively, these findings provide observational evidence in support of self-report data indicating that people control their emotions better with age (cf. Lawton et al., 1992). Although cross-sectional comparisons render conclusions about changes with age highly speculative, our profile of findings is also consistent with Guilford and Bengtson's (1979) longitudinal evidence for increasing positivity and decreasing negativity in marriage in late life. Placed within a broader life span model, these findings are consistent with socioemotional selectivity theory's (Carstensen, 1991, 1993) notions that emotion regulation assumes increasing importance with age and that people take active steps to avoid negative emotional experiences by negotiating the emotional course of social interactions. By recording specific affect and affect exchanges, we were able to show that, within the context of the marital relationship, older people use strategies that optimize emotional experiences and minimize negative emotional experiences.
Greater affection expressed by elderly couples during conflict resolution is consistent with other kinds of data we have obtained from these same couples. On questionnaires, elderly couples reported less disagreement than middle-aged couples across a wide range of issues (Levenson et al., 1993). In the domains of physiology and affective self-ratings, older couples, relative to middle-aged couples, evidenced lower levels of autonomic arousal, which we interpreted as being consistent with more positive emotional states (Levenson, Ekman, & Friesen, 1990), and rated their own affect as being more positive during conflict resolution (Levenson et al., 1993).

Inconsistent with Illig's (1977) general conclusion that there is greater predictability in the interaction styles of older couples, we found only one age difference for the five affect sequences that were studied (older unhappy couples were less likely to manifest negative start-up sequences than our other groups of couples). Thus, to the extent that affect sequences reflect temporal stability, our findings lend little support to this claim.

**Satisfaction**

Our finding of greater amounts of negative affect and negative continuance sequences in unhappy marriages than in happy marriages is consistent with a large literature (see Schaap et al., 1988, for a review), including our own previous studies with younger couples (Gottman, 1979; Levenson & Gottman, 1983, 1985). Moreover, the results of the present study generally suggest that this close relation between negative affect and marital unhappiness is maintained even in long-term marriages in middle age and old age. Although it might have been predicted that negative affective expression in unhappy marriages would diminish over the years (whether as a result of age or length of marriage), such that it no longer would be a distinguishing feature of long-term unhappy marriages, the present findings indicate that this is not the case. Even in marriages that have lasted more than 35 years, the link between marital unhappiness and expression of negative affect remains strong. However, our finding that older unhappy couples were less likely to engage in negative start-up sequences than other couples suggests that the former have achieved some control over the emergence of negative affect.

Without in any way diminishing the importance of global measures of emotion, one advantage of using a coding system such as SPAFF is its ability to move beyond global categories, such as positive and negative affect, to allow examination of the specific types of affect that characterize unhappy and happy marriages. In this sample, unhappy marriages were marked by a wide range of specific negative emotions including anger, contempt, sadness, and domineeringness.

The positive emotions of humor, affection, and validation were more likely to emerge in happy marriages than in unhappy marriages, even in a potentially hostile context such as discussing a marital conflict. Listener positivity and neutrality, often highly functional in helping to regulate conflictive interactions and in keeping the level of negative affect from escalating precipitously, were also more prevalent in happy marriages than in unhappy ones. Possibly indicative of this tighter control of negative affect, happy couples in this sample were less likely than unhappy couples to manifest sequences of negative affective continuance (negative affect followed by negative affect), a sequence that, if not regulated, can quickly lead to negative affect escalation.

**Gender**

In terms of emotional behavior, wives were clearly more emotionally expressive than husbands. Wives showed greater emotion overall, greater negative emotion, more anger, more contempt, more sadness, more whining, more joy, and more positive and negative affect as listeners. Also, they were more likely to engage in positive affect continuance. The two behaviors that husbands manifested at higher levels...
than wives were defensiveness (which may, in fact, represent emotional restraint or self-protection) and, in unhappy marriages, the affect sequence of deescalation (which may suggest a preference for nonemotional interaction). Taken as a whole, these results are consistent with a large literature on marital interaction in relatively young marriages that has found women to be more confronting and more affectively negative than men, who tend to be more defensive and more likely to try to escape from conflict (Christensen & Heavey, 1990; Gottman & Krokoff, 1989; Gottman & Levenson, 1988; Notarius & Johnson, 1982; Schaap, 1982).

In several areas, spousal differences were found only in unhappy marriages. For total speaker emotion, negative speaker emotion, contempt, and negative listener emotion, wives expressed more emotion than husbands only in unhappy couples. Thus, it appears that gender differences in negative emotion are particularly exacerbated in unhappy marriages.

In terms of notions that gender differences decrease with age (e.g., Gutmann, 1987; Hyde & Phillis, 1979; Keith & Brubaker, 1979), we found no interactions between gender and age. These findings, together with those of Zietlow and Sillars (1988), suggest that gender differences in marital interactive behavior are maintained in later stages of the life cycle. One caveat, however, is that our older couples are still relatively young. In addition, because we hope to examine marital interaction before and after retirement, our sample was limited to couples who had not yet retired. Thus, it certainly is possible that gender differences will lessen as our couples progress into later life, spending more time together (Gilford, 1984), coping with increasing frailty (especially of husbands; Troll & Bengtson, 1982), and moving through other late life transitions. We remain optimistic that observational research will provide an excellent modality by which to study gender differences. Unlike self-reports, which can be influenced by demand characteristics and other reporting biases, we expect that gendered behavior is less likely to be under the conscious awareness of the actor and thus less likely to be actively modulated.

Conclusion

This study has begun a process of applying observational methods of coding marital behavior to the interactional processes of a sample of marriages that are of much longer duration and whose spouses are much older than those that have typically been studied in the marital interaction literature. Both differences and similarities as a function of age were revealed. In terms of differences, important disparities between the emotional qualities of marriages in middle age and old age were found. Notably, in older marriages, the resolution of important conflicts was less negatively emotional and more affectionate than in middle-aged marriages. In terms of similarities, emotional differences between husbands and wives and between happy and unhappy marriages that have been found in other studies of younger couples were also found in these older couples. Most significant, the greater negative affect expressed in unhappy marriages than in happy marriages and the greater likelihood of wives to be affectively negative and husbands to be defensive and neutral were also found with these older cohorts.

Although the sampling procedure used to obtain these couples was designed to maximize generalizability of findings, it must be mentioned that the sample was drawn from one particular region of the country and that these data are cross sectional (thus confounding cohort and age differences). Ideally, to study the emotional differences between couples in their 40s and those in their 60s, one would conduct a 20-year longitudinal study. We hope to be able to do just this. There is a great deal more to learn about these couples as they change over time, and we hope to be able to continue following this sample as the middle-aged couples enter old age and as the older couples pass through retirement and beyond.

References


Gottman, J. M. (1989). *The Specific Affect Coding System (SPAFF).* (Unpublished manuscript, University of Washington, Seattle)


Illig, D. P. (1977). *Distributional structure, sequential structure, multivariate information analysis, and...*


Despite our efforts, we were unable to meet our recruitment goal for minority couples and subsequently oversampled White couples.

Our decision to separate the two age groups by 20 years (i.e., one generation) reflects our hope that we will be able to follow this sample for at least 20 years. At that time, the middle-aged couples will have reached the current age and marital duration of our older couples. This would afford us the opportunity to determine whether age-related differences found in the current cross-sectional analyses show evidence of being present longitudinally as well.

The $z$ scores were computed by taking the difference between the conditional probability of a given code following a given antecedent code (e.g., the probability that husband negative affect followed wife neutral affect) and the unconditional probability of the consequent code (in this example, the probability of husband negative affect) and dividing this difference by an estimate of the standard deviation (obtained with Allison and Liker's, 1982, formula). We have used similar procedures to quantify affect sequences in our previous marital research (e.g., Levenson & Gottman, 1985).

Table 1.
<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1/16</td>
<td>1/16</td>
<td>1/16</td>
<td>1/16</td>
<td>1/16</td>
</tr>
<tr>
<td>1/16</td>
<td>1/16</td>
<td>1/16</td>
<td>1/16</td>
<td>1/16</td>
</tr>
<tr>
<td>1/16</td>
<td>1/16</td>
<td>1/16</td>
<td>1/16</td>
<td>1/16</td>
</tr>
<tr>
<td>1/16</td>
<td>1/16</td>
<td>1/16</td>
<td>1/16</td>
<td>1/16</td>
</tr>
</tbody>
</table>